

XX

[illegible]

Db	121	tgagagctggagacgcygagtgcatgacgtgcgcatacgaaggggtccagggatgagatgccc	180
QY	181	TGTCCTAGATGTCAGAGCTGGAAACAAACAGAGAGACTGTGTGTGGTCTGGGAGAAATGT	240
Db	181	tgctctagaatgccaagctggaacaaacaaagaagacgtgttctgtgctctgggagatgt	240
QY	241	AATCATTCCTTCACCAACTGTCGATGCTCCCTGTGGGTGGAACGAACAATCCCTGCCCT	300
Db	241	aatcatctctccacacactgctgcatgctccctgtgggtggaacagaacaatgctgcacct	300
QY	301	CTCTGCCACGAGAGACTGGGTGTGCCAAGAAATGGCAATGAAGTGTAGAGGCTTC	360
Db	301	ctctgcgcacgagactggtgtgtgtccaaagaatcggcaaatgagatggtatgaagcttc	360
QY	361	TTAGGCGAGTGTCCAGAGCCCTGGTGATCTGTATCTCAGGCGCCTACAAAGGCTAG	420
Db	361	ttagcgagtggtccagagccctggtgtgatactgttaatccagtgccctaaagaagctaga	420
QY	421	ACACTACAGGGGATGAATTTCTTCAATAGAGGCCGATGATCTGTGGTCTTTGGACTCAT	480
Db	421	acactacagggatgaaatcttccaatagaagccgatgatactgtgtcttgactcat	480
QY	481	CAAGGCCCTGGTGTGCAATTTGTGCTTTTATCTTCAGAAATTCCTGTGATTTAGAAGAT	540
Db	481	caaaagccttggtatgcatctgttcaagtttcttccagaaatcttcgtgttaagaagat	540
QY	541	AATTTAATTAAGGTGTCTTCTTACCTCTGTGGTGTGTGCGGACACAGCTTTGAAG	600
Db	541	aatttaataaaggtgtccttccctaccctctgtgtgtgtgcgcacaagcttagaag	600
QY	601	TGCTATTAATAAAGGAAGAGCTCCAAATGAATACACTTATTTATACCATTTCTATAC	660
Db	601	tgcataaaaaaaggaagagatccaaatgaaatcaactataatcaaccatctcatac	660
QY	661	AACAGCAGTGAAGCAGATTTCAGACTTTTTCGATGCTTATGTTGATCAGTTAAAAAA	720
Db	661	aacagcgagtggaagcagtttcgagacttttcgatagtctatggtatcagtgataaanaa	720
QY	721	GAATGTTACAGTATCAAAATTAAGGCGAGTTTAA	754
Db	721	gaatgttaacagtaaccaataaagaatgcagttttaa	754

RESULT	4
AAx87320	
ID	AAx87320 standard; cDNA; 754 BP.
XX	
AC	AAx87320;
XX	
DT	27-SEP-1999 (first entry)
XX	
DE	Human sensitive to apoptosis (SAG) gene mutant MMA:
KX	
KW	SAG gene: sensitive to apoptosis; human; cancer; tumour;
KW	neurodegenerative disease; muscular dystrophy; wound healing;
KW	vulneray; therapy; mutant; ds.
XX	
OS	Homo sapiens.
OS	Synthetic.
XX	
FH	Location/Qualifiers
FT	1..342
FT	/*tag- a
FT	replace(181,T)
FT	/*tag- b
FT	/note- "C61S mutation"
XX	
PN	MO9932514-A2.
XX	
PD	01-JUL-1999.
XX	
15-DEC-1998;	98WO-US26705

XX 11-SEP-1998; 98US-0099840.
PR 19-DEC-1997; 97US-0068179.
XX
XX
PA (WARN) WARNER LAMBERT CO.
XX
XX
PI Sun Y;
XX
XX WPI, 1999-430152/36.
DR P-PSDB; AAY06498.
XX
XX
PT SAG: Sensitive to Apoptosis Gene and related proteins, useful for
PT promoting cell growth and protecting cells against apoptosis
XX
ES Claim 15, Page 64-65; 84pp; English.

CC This is the nucleotide sequence of human sensitive to apoptosis
CC (SAG) mutant gene MM4, which codes for a SAG protein (see AA006498)
CC in which the Cys residue at position 61 of the native protein (see
CC AA006492) is replaced by a Ser residue owing to a mutation of codon
CC 50 from TGC to AGC obtained by site-directed mutagenesis of SAG
CC cDNA. This residue is at zinc-finger finger 1 of SAG. Single and
CC double SAG mutants (see AA087117-31) were made in order to determine
CC the role of each cysteine residue of SAG in haem binding and SAG
CC oligomerization. These properties were unaffected by the MM4
CC mutation. SAG is a novel zinc finger protein that promotes
CC cell growth, protects cells from apoptosis, scavenges oxygen
CC radicals and can be used for the reversal of a tumour phenotype.
CC SAG genes, and mutant SAG genes, can be used to protect cells from
CC apoptosis induced by redox reagents. They can also be used for the
CC recombinant production of SAG proteins, which are molecular targets
CC in the development of drugs against neurodegenerative disorders,
CC cancers and muscle dystrophy, and promoting wound healing.

Query Match	99.8%	Score 752.4	DB 20	Length 754
Best Local Similarity	99.98%	Pred. No. 7.3e231		
Matches 753; Conservative	0	Mismatches 1	Indels 0	Gaps 0

QY	1	ATTCGCGACGTGGGAAGACGGAGAGAACTGGGGCCTTGGCCCTCTACTCTCCGGAGCTCA	60
Db	1	atggcgacagtggaagacggagagaaacctgcgcccctggccctcaactccggagagctca	60
QY	61	GGCCCAAGTCGGGAGGCGCAAGAATTTCTTCCCTCAAGAAATGTGAAACGGTGGGGCATG	120
Db	61	ggtcccaagtgcggagggcgacaagatgttctccctcaagaagtggaaacggyfgygcattg	120
QY	121	TGGAGCTGGGACGTGGAGTGGCATCTGCGCCATCTGCAAGGGTCCAGGTGATGATGCC	180
Db	121	tggagcctgggacgtggaggtgcgatacgtgcgcaatctgcagaggtccaggtgcagtgatgctc	180
QY	181	TGTTTGATGTCAAGCTGAAACAAACAAGAGACTGTTTGATGCTGGGAGAAATG	240
Db	181	agcttagatgctcaagtcgaagtgaataaacaagaagagactgtgtgtgtctctgggaagaatgt	240
QY	241	AATATTTCCTTCCACAACCTGCTGATCTCCCTTGGGTGAACAAGAAATCCCTGCCCT	300
Db	241	aatattctcttccacaactctgtcatctccctctgtgtgtgaacaagaacaatcgtctccct	300
QY	301	CTTCGCCAGAGGACTGGGTGGTCCAAAGAATGGGCAAAAGAGAGTGGTTAGAAGGCTTC	360
Db	301	ctctgcagcagaagcgtcgggctggcccaagaatctgcgaatggaagatggtttagaaggcttc	360
QY	361	TTAGCGCAGTTGTTTCAGAGCCGCTGTGGATCTTGTATCCAGTCCCTTACAAAGGCTAGA	420
Db	361	ttagcgagctgtgttcagagccctgtgtgattcttgtaatccagtgccctacaagaagctaga	420
QY	421	ACACTACAGGGGATGAATTCTTCAAAATAGAGCCGATGATCTGTGTCTTTTGACTCAT	480
Db	421	acatacaaggagatgaattcttcaaaatagaagacgcatgctgtgtgtcttcttggactcat	480

QY	481	CAAGCCCTGGTTAGCAATTTGTGAGTTTAACTTACAGAAATTCGTCGTAAAGAAAT	540
Db	481	caaaagcccttggttagcaatttgcagtttacttccagaaattcccttgatlaagaagat	540
QY	541	AATTTAATTTAAAGGCGATCTTCTTACCTCTGTGGTGTGTGCGGCACACAGCTTAGAG	600
Db	541	aattatttaaagggcgctcccttcactctcgtggtggtgcgcacacaagcttgaag	600
QY	601	TGCTATATAAAAAGAGAGAGCTCCAAATTGATCAGCTTATTAATTACCATTTCTATAC	660
Db	601	tgctataaaaaaagaaagagctcccaaatgcatcacccttaattaccattctctac	660
QY	661	AACAGCAGAGGAGGAGCGAGTTTTCGAGCTTTTTCGATGCTTATGTTGATCAGTTAAAAA	720
Db	661	aacagcagagtggaagcagcttcgcagacttctgcagtcatggttgatcagttaaaaa	720
QY	721	GAATGTTACAGTAACAATAAAGCGAGTTTAAA	754
Db	721	gaatgttacagtaacaataaagcgagttttaa	754
RESULT	5		
ID	AAx87321		
XX	AAx87321	standard; cDNA; 754 BP.	
AC	AAx87321;		
DT	27-SEP-1999	(first entry)	
XX			
DE		Human sensitive to apoptosis (SAG) gene mutant MM5.	
XX			
KW		SAG gene; sensitive to apoptosis; human; cancer; tumour;	
KW		neurodegenerative disease; muscular dystrophy; wound healing;	
KW		vulnerary; therapy; mutant; ds.	
XX			
OS		Homo sapiens.	
OS		Synthetic.	
XX			
FH	Key	Location/Qualifiers	
FT	CDS	1..342	
FT		/*tag= a	
FT	mutation	replace(238,T)	
FT		/*tag= b	
FT		/note= "C80S mutation"	
XX			
PN	WO932514-A2.		
XX			
PD	01-JUL-1999.		
XX			
PF	15-DEC-1998;	98WO-US26705.	
XX			
PR	11-SEP-1998;	98US-0099840.	
PR	19-DEC-1997;	97US-0068179.	
XX			
PA	(WARN) WARNER LAMBERT CO.		
XX			
PI	Sun Y;		
XX			
DR	WPI: 1999-430152/36.		
DR	P-PSDB; AAY06499.		
PT	SAG: Sensitive to Apoptosis Gene and related proteins, useful for		
XX	promoting cell growth and protecting cells against apoptosis		
XX	Claim 15; Page 66; 84pp; English.		
XX			
CC	This is the nucleotide sequence of human sensitive to apoptosis		
CC	(SAG) mutant gene MM5, which codes for a SAG protein (see AAY06499)		
CC	in which the Cys residue at position 80 of the native protein (see		
CC	AAY06492) is replaced by a Ser residue owing to a mutation of codon		
CC	80 from TGC to AGC obtained by site-directed mutagenesis of SAG		
CC	cDNA. This residue is in zinc-fing finger 1 of SAG. Single and		
CC	double SAG mutants (see AAX87317-31) were made in order to determine		

CC the role of each cysteine residue of SAG in haem binding and SAG
CC oligomerization. These properties were unaffected by the MMS
CC mutation. SAG is a novel zinc finger protein that promotes
CC cell growth, protects cells from apoptosis, scavenges oxygen
CC radicals and can be used for the reversal of a tumour phenotype.
CC SAG genes, and mutant SAG genes, can be used to protect cells from
CC apoptosis induced by redox reagents. They can also be used for the
CC recombinant production of SAG proteins, which are molecular targets
CC in the development of drugs against neurodegenerative disorders,
CC cancers and muscle dystrophy, and promoting wound healing.
XX
SQ Sequence 754 BP; 206 A; 155 C; 201 G; 192 T; 0 other;

Query Match	Similarity	99.8%	Score	752.4	DB	20	Length	754
Best Local	Similarity	99.9%	Pred.	No.	7.3e-231			
Matches	753	Conservative	0	Mismatches	1	Indels	0	Gaps
QY	1	ATGCGCGAGCTGGAAAGACGAGAGAGAAACCTTGCSCCTTGCCCTCTCACTCCGGGAGCTCA	60					
Db	1	atgycgcagcgttgyaagacgagagaagaaacctgcgccttgcctctcaactccggagactca	60					
QY	61	GGCTCCAGTGGGGGAGCGACAGAAAGTGTCTCCCTCCAAAGAACTGGAAACCGCGGTGGCGAAG	120					
Db	61	ggctccaagatcgggagcgacaaagatgtctctccccaagaagtggaacgcggtgcgacag	120					
QY	121	TGGAGCTGGGAGCGTGGAGATGCGATTCGTCGGCCATCTGCAGGGCTCCAGGTGATGATGTC	180					
Db	121	tggagctgggagcgatgagatgcgatacgtgcgcatactggaaggtcccaagtgatgtatgac	180					
QY	181	TGTCTTAGATGTCGAAGCTTAAAAACAACAAGAGACTGTGTGTGTGTCTGGGAGAAATCT	240					
Db	181	tgctcttagatgtgtcaagcttgaataaacaacaagaagactgtgtgtgtcttggggagaagat	240					
QY	241	AATCATTCCTTCCACAAACGCTGCATGTCTCCCTGGGTGAACAGAACAAATGCTCCCT	300					
Db	241	aatcatctcttccacaacactgcctgcatagtctccctgtgtgtgtgaacaagacaatactgcct	300					
QY	301	CTGTGCGACGAGACTGGGTGTCCAAAGATTCGCAAAATGCGCAATGAGAGTGTTCAGAGCTTC	360					
Db	301	ctgtgcgacgagactgggtgtgtccaaagatgtcgaatgcgcaaatgagagtgtgaagaagcttc	360					
QY	361	TTAGGGCAATGTCAGACCCCTGGTGGATGTCGTAAATCAGAGCCCTCAAAAGCTAGA	420					
Db	361	ttagggcgaatgtctcagacccctgggtggatgtcgtaaatcagagccctcaaaagctaga	420					
QY	421	ACACTACAGAGGANTGAATCTTCAATATAGAGCCGATGATGTGTGCTTTGGACTCAT	480					
Db	421	acaactacagggatgaattcttcaatatagagccgagtgtgactgtgtcttggactcat	480					
QY	481	CAAAGCCTTGTTAGCATTTGTACAGTTTATCTTCAGAAATCTCTGTATTAAGAAGAT	540					
Db	481	caaaagccttgttagcatattgtcagtttatacttcgaataatctctgtatlaagaagat	540					
QY	541	AATTTATTTAAAGTGTGCTCTTCTACCTCTGTGAGTGTGTGGCGACACAGCTTTGAG	600					
Db	541	aatttaattaagtggtgtcttcttaccttacctctgtgtgtgtgtgcgcacacagcttaag	600					
QY	601	TGCTATAAAAAAGAAAGAGCTCCAAATTTGAATCACCTTAATAATTTACCATTTCTAAC	660					
Db	601	tgcataaaaaaagaaagagctccaattgaatcacctataattaccatttccataac	660					
QY	661	AACAGGCAATGGAAGCATTTTGAGACTTTTTCGAGCTTAATGAGTATGATCAGTTAAAAAA	720					
Db	661	aacaggcaatggaaagcatgttgcagacttcttcgaigtcttaatgtgtatcagttataaaaa	720					
QY	721	GATGTTACAGTACAAATTAAGTGCAGTTTAAA	754					
Db	721	gatgttacagtacaataataaagtgtgagttaa	754					

ID	AAx87324	strand; cDNA; 754 BP.
AC	AAx87324;	
XX		
DT	27-SEP-1999	(first entry)
XX		
DE	Human sensitive to apoptosis (SAG) gene mutant MM8.	
XX		
KW	SAG gene; sensitive to apoptosis; human; cancer; tumour;	
KW	neurodegenerative disease; muscular dystrophy; wound healing;	
KW	vulnerary; therapy; mutant; ds.	
XX		
OS	homo sapiens.	
OS	Synthetic.	
XX		
FM	Key	Location/Qualifiers
FT	CDS	1..342
FT	mutation	/*tag= a replace(262,T) /*tag= b /note= "C885 mutation"
FT		
FT		
XX	WO9932514-A2.	
XX		
XX	01-JUL-1999.	
XX		
PD	15-DEC-1998;	98MO-US26705.
XX		
PF	11-SEP-1998;	98US-0099840.
PR	19-DEC-1997;	97US-0068179.
XX		
XX	(WARN)	WARNER LAMBERT CO.
PA		
XX		
XX	Sun Y;	
PI	WPT; 1999-430152/36.	
XX		
DR	P-PSDB; AAY06502.	
XX		
XX	SAG: Sensitive to Apoptosis Gene and related proteins, useful for	
PT	promoting cell growth and protecting cells against apoptosis	
XX		
XX	Claim 15; Page 71; 84pp; English.	
PS		
XX		
XX	This is the nucleotide sequence of human sensitive to apoptosis	
CC	(SAG) mutant gene MM8, which codes for a SAG protein (see AAY06502)	
CC	in which the Cys residue at position 88 of the native protein (see	
CC	AAV06492) is replaced by a Ser residue owing to a mutation of codon	
CC	88 from TGC to AGC obtained by site-directed mutagenesis of SAG	
CC	cDNA. This residue is in zinc-finger finger 2 of SAG. Single and	
CC	double SAG mutants (see AAX87317-31) were made in order to determine	
CC	the role of each cysteine residue of SAG in haem binding and SAG	
CC	oligomerization. These properties were unaffected by the MM8	
CC	mutation. SAG is a novel zinc finger protein that promotes	
CC	cell growth, protects cells from apoptosis, scavenges oxygen	
CC	radicals and can be used for the reversion of a tumour phenotype.	
CC	SAG genes, and mutant SAG genes, can be used to protect cells from	
CC	apoptosis induced by redox reagents. They can also be used for the	
CC	recombinant production of SAG proteins, which are molecular targets	
CC	in the development of drugs against neurodegenerative disorders,	
CC	cancers and muscle dystrophy, and promoting wound healing.	
XX		
XX	Sequence 754 BP; 206 A; 155 C; 201 G; 192 T; 0 other;	

Query Match	99.8%;	Score 752.4;	DB 20;	Length 754;
Best Local Similarity	99.9%;	Pred. No. 7.3e-231;		
Matches 753; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0

0y 1 ATGGCGCAGCGTGAAGAGAGAGAGAAACCTGCGCCTTCACTCCGGAGCTCA 60
|||||
Db 1 atggcgcagcgtggaagagagagaagaaacctgcgccttcaactccggaagtcca 60
61 ggcgtccagtcgtgggagggcgacaaagtgtttcccttcagaagaagtggaaacgcggtggccatg 120

Db	61	ggtcccaagctcgggagggcgaagaagtgtctccctccaagaagtggaacgcggtggccatg	1200
QY	121	tggagcttgsgagctgaggatgacgtacgtccgccaattgccaaggtccagtgatgagatgcc	1800
Db	121	tggagcttgsgagctggaatgcatagctgctccatctccgaggtgccaggtgcatgatagcc	1800
QY	181	tgtcttagatgtcgaagcttgaaaaacaaagaagactgtgtgtgtcctggggagaaatgt	2400
Db	181	tgtcttagatgtcgaagcttgaaaaacaaagaagactgtgtgtgtcctggggagaaatgt	2400
QY	241	aattcatttccttccacaactgscatgtgcctctgggggtgaaacagacaaatccgtgcctt	3000
Db	241	aattcatttccttccacaactgscatgtgcctctgggggtgaaacagacaaatccgtgcctt	3000
QY	301	ctctccagcagagactgggtgtccaaagaatccgcaaatgagatggtttagaagagcttc	3600
Db	301	ctctcgccagcagactgggtgtccaaagaatccgcaaatgagatggtttagaagagcttc	3600
QY	361	tttagcgcgagttgttcagagcccttgatgtttaaTCCAGTGCCTACCAAAAGCTAGA	4200
Db	361	tttagcgcgagttgttcagagcccttgatgtttaaTCCAGTGCCTACCAAAAGCTAGA	4200
QY	421	acacacagcagggatcaattcttcaaatgagccgagatcctgtgcttTGGAGCTCAT	4800
Db	421	acacacacagggatgaaatcttccaaatgagccgagatcctgtgcttTGGAGCTCAT	4800
QY	481	caaaagccttggtttagcatttgcagttttatcttCAGAAATCTCTGTGATTAAGAAGAT	5400
Db	481	caaaagccttggtttagcatttgcagttttatcttCAGAAATCTCTGTGATTAAGAAGAT	5400
QY	541	aatttatTAAGGTGTCCTTCCACTCTGTGTGTGTGTGCGGCACACAGCTTAGAAG	6000
Db	541	aatttatTAAGGTGTCCTTCCACTCTGTGTGTGTGTGCGGCACACAGCTTAGAAG	6000
QY	601	tgcTATAAAAAAGAAAGAGCTCCAAATTGATACCTTTAAATTATCCCATTTCTATAC	6600
Db	601	tgcTATAAAAAAGAAAGAGCTCCAAATTGATACCTTTAAATTATCCCATTTCTATAC	6600
QY	661	aaacagcagtggaagcagtttcgagacttttcgactgtatgattgattcagttaaaaaa	7200
Db	661	aaacagcagtggaagcagtttcgagacttttcgactgtatgattgattcagttaaaaaa	7200
QY	721	gaatgtTACGATACAAATAAAGGCGAGTTAAA	754
Db	721	gaatgtTACGATACAAATAAAGGCGAGTTAAA	754

	RESULT	7
AA87325		
ID	AA87325	standard; cDNA; 754 BP.
XX		
AC	AA87325;	
XX		
DT	27-SEP-1999	(first entry)
XX		
DE	Human sensitive to apoptosis (SAG) gene	
XX		
KW	SAG gene; sensitive to apoptosis; human	
KM	neurodegenerative disease; muscular dy-	
KW	vulnerable; therapy; mutant; ds.	
XX		
OS	Homo sapiens.	
OS	Synthetic.	
XX		
PH	Key	Location/Qualifiers
FT	CDS	1..342
FT		/*tag= a
FT	mutation	replace(295,'T')
FT		/*tag= b
FT	/note= "C99S mutation	
XX		
PN	W09932514 -A2.	

XX 01-JUL-1999.
 PD 15-DEC-1998; 98WO-US26705.
 PF 11-SEP-1998; 98US-0099840.
 PR 19-DEC-1997; 97US-0068179.
 XX (WARN) WARNER LAMBERT CO.
 PA Sun Y;
 PI WPI; 1999-430152/36.
 DR P-PSDB; AAY06503.
 XX
 XX SAG: Sensitive to Apoptosis Gene and related proteins, useful for
 PT promoting cell growth and protecting cells against apoptosis
 PS Claim 15; Page 72-73; 84pp; English.

CC This is the nucleotide sequence of human sensitive to apoptosis
 CC (SAG) mutant gene MM9, which codes for a SAG protein (see AAY06503)
 CC in which the Cys residue at position 99 of the native protein (see
 CC AAY06492) is replaced by a Ser residue owing to a mutation of codon
 CC 99 from TGC to AGC obtained by site-directed mutagenesis of SAG
 CC cDNA. This residue is in zinc-finger 2 of SAG. Single and
 CC double SAG mutants (see AAX87317-31) were made in order to determine
 CC the role of each cysteine residue of SAG in haem binding and SAG
 CC oligomerization. These properties were unaffected by the MM9
 CC mutation. SAG is a novel zinc finger protein that promotes
 CC cell growth, protects cells from apoptosis, scavenges oxygen
 CC radicals and can be used for the reversion of a tumour phenotype.
 CC SAG genes, and mutant SAG genes, can be used to protect cells from
 CC apoptosis induced by redox reagents. They can also be used for the
 CC recombinant production of SAG proteins, which are also used for the
 CC in the development of drugs against neurodegenerative disorders,
 CC cancers and muscle dystrophy, and promoting wound healing.

Sequence 754 BP; 206 A; 155 C; 201 G; 192 T; 0 other;

Query Match 99.8%; Score 752.4; DB 20; Length 754;
 Best Local Similarity 99.9%; Pred. No. 7.3e-231;
 Matches 753; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCCGAGCTGGAGAGCGAGAAACCTGCGCCCTCTACTCCGGAGACTA 60
 DB 1 atggccgagctggagagcgagaaacctgcccctctcactccgagactca 60
 QY 61 GGCTCCAGTGGGAGGAGGACAAAGATGTCCTCCCAAGAGGAAAGCGGGCCATG 120
 DB 61 ggctccagtgaggagagcgacaaagatgtctccccaagagtgagagcgagccatg 120
 QY 121 TGGAGCTGGAGCTGGAGTGCATACGTGCGCCATCTGAGGGTCCAGTGTATGATGCC 180
 DB 121 tggagctggagctggagtgcatatcgctgcccattctgcaaggttccagtgatgagtc 180
 QY 181 TGTCTTAGATGTCAAGCTAAAAACAAGAGAGTGTGTGTGTGTGTGGGAGAAATGT 240
 DB 181 tgtcttagatgtcaagctaaaaacaagagagtggtgtgtgtgtgtgtgtgtgtgtgtgt 240
 QY 241 AATCATTCCTTCACAACTGCTGATGCTCCCTGTGGGTAAACAGAAACAATGCTCCCT 300
 DB 241 aatcatctcttcacaaactgctgatgctccctgtgggttaaacagaaacaatgctccct 300
 QY 301 CTCTGCCACGAGACTGGTGTGTCGAAGATGCGCAAAATGAGAGTGTGTAGAGGCTTC 360
 DB 301 ctctgccacgagactggtgtgtcgaaagatgcgcaaaatgagagtgtgttagagcttc 360
 QY 361 TTAGCGCAGTGTGTGAGAGCCCTGGTGGATCTTGTATACAGGCCCTCAAGAGCTAGA 420
 DB 361 tttagcgagtggtgtgagagccctgggtggatcttgtatcagagccctcaagagctaga 420

QY 421 ACACTACAGGAGATGAATTCCTCAATAGAGCCGATGATCTGTGCTTTGGACTCAT 480
 DB 421 acactacagggatgatacttctcaaatagagccgatgatactgtgtcttggactcat 480
 QY 481 CAAAGCCTTGTTAGCATTTGTACAGTTTATCTTCAGAAATCTCTGTATTAAGAAGAT 540
 DB 481 caaagccttgtagcatcttgcagtttatacttcagaaatctctgtatgaagaagat 540
 QY 541 AATTTTAAAGTGTGCTTCTTACCTGTCGTCGTGTGTGTGTGTGTGTGTGTGTGTGT 600
 DB 541 aattttaagtggtgcttcttacctgtcgtcgtgtgtgtgtgtgtgtgtgtgtgtgt 600
 QY 601 TCTATATAAAGAGAGAGCTCCAAATTAATCACTTATATTAATTAATTAATTAATTAAT 660
 DB 601 tctataaag 660
 QY 661 AACAGGCACTGGAAGCAGTTTGCAGACTTTTTCGATGCTTATGATGATGATGATGAT 720
 DB 661 aacaggcagtggaagcag 720
 QY 721 GAATGTTACAGTACAAATTAAGTGCAGTTTAA 754
 DB 721 gaatgttacagtacaaatgaagtgcatgatttaa 754

RESULT 8

AAX87327 standard; cDNA; 754 BP.

AAX87327;

27-SEP-1999 (first entry)

Human sensitive to apoptosis (SAG) gene mutant MM1.

SAG gene; sensitive to apoptosis; human; cancer; tumour;

neurodegenerative disease; muscular dystrophy; wound healing;

vulnerary; therapy; mutant; ds.

Homo sapiens.

Synthetic.

Key

mutation

W09932514-A2.

01-JUL-1999.

15-DEC-1998; 98WO-US26705.

11-SEP-1998; 98US-0099840.

19-DEC-1997; 97US-0068179.

(WARN) WARNER LAMBERT CO.

Sun Y;

WPI; 1999-430152/36.

P-PSDB; AAY06505.

Claim 15; Page 76; 84pp; English.

This is the nucleotide sequence of human sensitive to apoptosis
 CC (SAG) mutant gene MM1, which codes for a SAG protein (see AAY06505)
 CC in which the Cys residue at position 64 of the native protein (see

CC AAY06492) is replaced by a Ser residue owing to a mutation of codon
 CC 64 from TGT to AGC obtained by site-directed mutagenesis of SAG
 CC cDNA. This residue is in zinc-finger finger 1 of SAG. Single and
 CC double SAG mutants (see AAX87317-31) were made in order to determine
 CC the role of each cysteine residue of SAG in haem binding and SAG
 CC oligomerization. These properties were unaffected by the MM12
 CC mutation. SAG is a novel zinc finger protein that promotes
 CC cell growth, protects cells from apoptosis, scavenges oxygen
 CC radicals and can be used for the reversion of a tumour phenotype.
 CC SAG genes, and mutant SAG genes, can be used to protect cells from
 CC apoptosis induced by redox reagents. They can also be used for the
 CC recombinant production of SAG proteins, which are molecular targets
 CC in the development of drugs against neurodegenerative disorders,
 CC cancers and muscle dystrophy, and promoting wound healing.
 CC
 XX
 XX
 S0 Sequence 754 BP; 206 A; 155 C; 201 G; 192 T; 0 other;

Query Match 99.8%; Score 752.4; DB 20; Length 754;
 Best Local Similarity 99.9%; Pred. No. 7.3e-231;
 Matches 753; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCCGACGTGGAAGACGAGAGAAACCTGGCCCTGCTCTCACTCCGGAGCTCA 60
 Db 1 atggccgacgtggaagacgagagaaacctggccctgctctcaactccggagctca 60
 QY 61 GGCTCCAGTGGGAGGCGACAGATGTTCTCCCTCAAGAGTGAACGGCGTGGCCATG 120
 Db 61 ggctccagtgaggagcgacaagaatgtctccctcaagaagtgaacggcggtggccatg 120
 QY 121 TGGAGCTGGGACCTGGAGTGGATACGTGGCCATCTGCAGGGTCCAGGTGATGATGCC 180
 Db 121 tggagctgggacctggagtggtatgctggccatctgcaggggtccaggtgatgattgcc 180
 QY 181 TGTCTTAGATGTCAAGCTGAAACAACAAGAGACTGTGTGTGGTGGGGAAGTATG 240
 Db 181 tgtcttagatgtcaagctgaaacaacaagagactgtgtgtgggtggggaagtatg 240
 QY 241 AATCATCTCTTCCACACTGCTGCATGCTCCCTGGGTGAACAACAATCGCTGCCCT 300
 Db 241 aatcatctcttccacactgctgcattccctgggtgaaacaacaatcgctgccct 300
 QY 301 CTGCGCAGACGACGTGGGTGGTCCAAAGATGGCAAAATGACAGATGTTAGAAAGCTTC 360
 Db 301 ctgcgacgacgacgtgggtgggtccaaagatggcaaaatgacagatgttagaagcttc 360
 QY 361 TTAGCGCAGTTGTTACAGAGCCCTGGTGATCTTGAATCCAGTCCCTACAAAGGCTAGA 420
 Db 361 tttagcgagttgttacagagccctggtgatcttgaatccagtccctacaaaggctaga 420
 QY 421 ACACATACAGGGGATGAATTTTCAATATGAGACCCGATGATGATGATGTTGGACTCAT 480
 Db 421 acctataagggatgaattttcaatatgagacccgatgatgatgattggactcat 480
 QY 481 CAAAGCCTTGTTAGCATTTGTGAGTTTATCTTCAGAAATTCCTGCTGATTAAGAAGAT 540
 Db 481 caaaccttgttagcatttgtgagtttattcttcagaaattcctggtattaagaagat 540
 QY 541 AATTTATTAAGGTGCTCTTCTTACCTCTGTTGGTGTGTGTCGACACAGCTTGAAG 600
 Db 541 aatttattaaggtgctcttcttacctctggtgtgtgtgtgctgcgaacagcttgaag 600
 QY 601 TGTCTAATAAAGGAGAAAGAGCTCCAAATGGAATCACCTTAATAATTAACCATTTATAC 660
 Db 601 tgtctataaaggagaaagagctccaaatggaatcaccttaataattaccatttcatc 660
 QY 661 AACAGCAGTGAAGAGAGCTTTCAGACTTTTTCAGACTTTTTCAGACTTTTTCAGACTTT 720
 Db 661 aacagcagtgaagagagctttcagactttttcagactttttcagactttttcagacttt 720
 QY 721 GAATGTTACAGTAAACAATTAAGTGCAGTTTAAA 754
 Db 721 gaatgttacagtaacaataaagtgcagtttaa 754

RESULT 9
 ID AAX87328 standard; cDNA; 754 BP.
 XX
 AC AAX87328;
 XX
 DT 27-SEP-1999 (first entry)
 DE Human sensitive to apoptosis (SAG) gene mutant MM12.
 XX
 KW SAG gene; sensitive to apoptosis; human; cancer; tumour;
 KW neurodegenerative disease; muscular dystrophy; wound healing;
 KW vulnerability; therapy; mutant; ds.
 XX
 OS Homo sapiens.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT CDS 1..342
 FT mutation /tag= a
 FT replace(217,T)
 FT /tag= b
 FT /note= "C73S mutation"
 PN MO9932514-A2.
 PD 01-JUL-1999.
 XX
 PF 15-DEC-1998; 98MO-US26705.
 XX
 PR 11-SEP-1998; 98US-009840.
 PR 19-DEC-1997; 97US-0068179.
 XX
 PA (WARN) WARNER LAMBERT CO.
 XX
 PI Sun Y;
 XX
 DR WPI; 1999-430152/36.
 DR P-PSDB; AAY06506.
 XX
 PS Claim 15; Page 77-78; 84pp; English.

CC This is the nucleotide sequence of human sensitive to apoptosis
 CC (SAG) mutant gene MM12, which codes for a SAG protein (see AAY06506)
 CC in which the Cys residue at position 73 of the native protein (see
 CC AAY06492) is replaced by a Ser residue owing to a mutation of codon
 CC 73 from TGT to AGC obtained by site-directed mutagenesis of SAG
 CC cDNA. This residue is in a protease inhibitor motif of SAG. Single
 CC and double SAG mutants (see AAX87317-31) were made to determine
 CC the role of each cysteine residue of SAG in haem binding and SAG
 CC oligomerization. These properties were unaffected by the MM12
 CC mutation. SAG is a novel zinc finger protein that promotes
 CC cell growth, protects cells from apoptosis, scavenges oxygen
 CC radicals and can be used for the reversion of a tumour phenotype.
 CC SAG genes, and mutant SAG genes, can be used to protect cells from
 CC apoptosis induced by redox reagents. They can also be used for the
 CC recombinant production of SAG proteins, which are molecular targets
 CC in the development of drugs against neurodegenerative disorders,
 CC cancers and muscle dystrophy, and promoting wound healing.
 CC
 XX
 XX
 S0 Sequence 754 BP; 206 A; 155 C; 201 G; 192 T; 0 other;

Query Match 99.8%; Score 752.4; DB 20; Length 754;
 Best Local Similarity 99.9%; Pred. No. 7.3e-231;
 Matches 753; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCCGACGTGGAAGACGAGAGAAACCTGGCCCTGCTCTCACTCCGGAGCTCA 60

```

Db      1 atggccgaagctgtaagaacgagagaacacgtgcgcctcctcctcaccgagagctca 60
QY      61 GGCTCCAATCGGAGGCGACAAGATGTTCTCCCTCAGAAGTGCAGCCGGTGCCATG 120
Db      61 ggctccaagtgcggagcgacaagatgcttcctccaagaagtgaacgcgctgagcatg 120
QY      121 TGGAGCTGGAGGTGGAGTACGTACGTGCGCATCTCGAGGTCCAGGTGATGATGTCG 180
Db      121 tggagctggagctggagtgagtgacgacgacgacatcgcaggtccaggtgagtgcc 180
QY      181 TGTCTTAGATGTCAAGCTGAAAAAACAAGAGAGACTGTGTGTGCTGGGAGAGATGT 240
Db      181 tgtcttagatgtcaagctgtaaaacaagaagagacagtgctgtgctgggagagatgt 240
QY      241 AATCATTCCTTCACAACTGCTGCATGTCCTGCTGGGTGAAACAGCAATGCGTCCCT 300
Db      241 aatcatctcttcacaactgctgcatagtccctgtggtgaaacagacaatcgtccct 300
QY      301 CTCTGCCACGAGACTGGGTGGTCCAAAGATCGGCAATGAGAGTGTAGAGGCTTC 360
Db      301 ctctgccagcagactgggtggtccaaagaatcgcgcaaatgagatgtagaaggcttc 360
QY      361 TTAGCGCAGTTGTTACAGAGCCCTGTGTGATCTTGAATCAGTGCCTCAAAAGCTAGA 420
Db      361 tttagcgcaagtgttcagagccctgtgtgacatcctgacagtgccctacaagctaga 420
QY      421 AACACACAGGAGATGAATCTTCAAAATAGAGCCGATGATGTGTGCTTGGACTCAT 480
Db      421 aaacacagggagatgaatcttcaaaatagagcgatgaltgtgtcttgactcatc 480
QY      481 CAAGACCTTGTTACATTTGACATTTGATTTATTCAGAAATCTCTGATTAAGAGAT 540
Db      481 caagacctgtgtacattgttcagattatcttcagaatctctgtgataagagct 540
QY      541 AATTTAATTAAGGTGCTTCTTCACTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
Db      541 aatttataaagtgctctctcctcctcctcctcctcctcctcctcctcctcctcctc 600
QY      601 TGCTTTAAAAAAGGAAGAGCTCCAAATGATCACTTATTAATTAATTAATTAATTAAT 660
Db      601 tgccttataaaaggaagagctccaaatgataatcattataatcattcctatc 660
QY      661 AACAGGACAGTGGAGACAGTTTGCAGACTTTTGCATGTTGATGATCAGTTAAAAA 720
Db      661 aacaggacagtggaagcaggttcgagacttttcgattgctttagtgcagttaaaaa 720
QY      721 GAATGTTACAGTACAAATAAAGTGCAGTTTAAA 754
Db      721 gaatgttacagtaacaataaagtcagttaaa 754

```

RESULT 10

AAK87331
ID AAK87331 standard; cDNA; 754 BP.

AAK87331;

27-SEP-1999 (first entry)

Human sensitive to apoptosis (SAG) gene mutant MM15.

SAG gene; sensitive to apoptosis; human; cancer; tumour;

neurodegenerative disease; muscular dystrophy; wound healing;

vulnerary; therapy; mutant; ds.

Homo sapiens.

Synthetic.

Key CDS

FT mutation

Location/Qualifiers
1..342
/*tag=
replace(139,C)

```

FT      /*tag= b
FT      /note= "C47S mutation"
PN      W09932514-A2.
XX      01-JUL-1999.
PD      15-DEC-1998; 98WO-US26705.
PF      11-SEP-1998; 98US-0099840.
PR      19-DEC-1997; 97US-0068179.
XX      (WARN ) WARNER LAMBERT CO.
PI      Sun Y;
DR      WPI: 1999-430152/36.
DR      P-PSDB; AAY06509.
PT      SAG: Sensitive to Apoptosis Gene and related proteins, useful for
PS      promoting cell growth and protecting cells against apoptosis
PS      Claim 15; Page 82-83; 84pp; English.

```

This is the nucleotide sequence of human sensitive to apoptosis (SAG) mutant gene MM15, which codes for a SAG protein (see AAY06509) in which the Cys residue at position 47 of the native protein (see AAY06492) is replaced by a Ser residue owing to a mutation of codon 47 from TGC to AGC obtained by site-directed mutagenesis of SAG cDNA. This residue is in a GADPH binding site of SAG. Single CC and double SAG mutants (see AAK87317-31) were made to determine the role of each cysteine residue of SAG in haem binding and SAG oligomerization. These properties were unaffected by the MM15 mutation. SAG is a novel zinc finger protein that promotes cell growth, protects cells from apoptosis, scavenges oxygen radicals and can be used for the reversal of a tumour phenotype. CC SAG genes, and mutant SAG genes, can be used to protect cells from apoptosis induced by redox reagents. They can also be used for the recombinant production of SAG proteins, which are molecular targets in the development of drugs against neurodegenerative disorders, cancers and muscle dystrophy, and promoting wound healing.

Sequence 754 BP; 206 A; 155 C; 201 G; 192 T; 0 other;

Query Match

99.8%; Score 752.4; DB 20; Length 754;

Best Local Similarity 99.9%; Pred. No. 7.3e-231;

Matches 753; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY      1 ATGGCCGACGTGGAGAGCAGAGAGAAACCTGCGCCTGGCTTCACATCCGGAGACTCA 60
Db      1 atggccgaagctgtaagaacgagagaacacgtgcgcctcctcctcaccgagagctca 60
QY      61 GGCTCCAATCGGAGGCGACAAGATGTTCTCCCTCAGAAGTGCAGCCGGTGCCATG 120
Db      61 ggctccaagtgcggagcgacaagatgcttcctccaagaagtgaacgcgctgagcatg 120
QY      121 TGGAGCTGGAGGTGGAGTACGTACGTGCGCATCTCGAGGTCCAGGTGATGATGTCG 180
Db      121 tggagctggagctggagtgagtgacgacgacatcgcaggtccaggtgagtgcc 180
QY      181 TGTCTTAGATGTCAAGCTGAAAAAACAAGAGAGACTGTGTGTGCTGGGAGAGATGT 240
Db      181 tgtcttagatgtcaagctgtaaaacaagaagagacagtgctgtgctgggagagatgt 240
QY      241 AATCATTCCTTCACAACTGCTGCATGTCCTGCTGGGTGAAACAGCAATGCGTCCCT 300
Db      241 aatcatctcttcacaactgctgcatagtccctgtggtgaaacagacaatcgtccct 300
QY      301 CTCTGCCACGAGACTGGGTGGTCCAAAGATCGGCAATGAGAGTGTAGAGGCTTC 360
Db      301 ctctgccagcagactgggtggtccaaagaatcgcgcaaatgagatgtagaaggcttc 360

```

QY 361 TTAGCGCAGTTGTTACAGACCCCTGGTGAGTCTGTAATCCAGTGGCCCTACAAAAGCTAGA 420
| | | | |
DB 361 tttagcgagttgttcacagagccctggtgaltctgttaaccagtgacctacaagaagctaga 420
QY 421 ACCTCAGGGGAGTATTTTCAATAGAGCCGATGATCTGTGCTTTGGACTCAT 480
| | | | |
DB 421 acctacaaggagatgaattctcaataagagccgagatcgtgtcttgactcat 480
QY 481 CAAAGCCCTTGTTAGCATTTGTGAGTTTATCTTACAGAAATTCCTGTGATTAAGAAGAT 540
| | | | |
DB 481 caaagccttgtagtacctgttcagtttcttcctcgaattctctgtgttaagaagat 540
QY 541 AATTATTAAGGTGCTCTCTACCTCTGTGGTGTGTCTGGCGACACAGCTTAGAAG 600
| | | | |
DB 541 aattatlaaaggtygtccctcctaccctcgtgtgtgtcgcgcacacagcttagaag 600
QY 601 TGGTATTAAGAAAGAAAGAGCTCCAAATTAATCACCCTTAAATTTACCATTTCTATAC 660
| | | | |
DB 601 tgcataaaaaaagaaagagctccaattgatacccttaatttaccattctctatc 660
QY 661 AACAGCAGTGGAAAGCAGTTTGCAGACTTTTGCATGCTTATGTTGATGAGTTAAAAA 720
| | | | |
DB 661 aacagcgacgtggagagcagcttgcagacttcttgatgcttatagttgatcagttaaaaa 720
QY 721 GAATGTTACAGTAACAATAAAGTCAGTTTAAA 754
| | | | |
DB 721 gaatgtacagtaacaataaagtcagtttaa 754

RESULT 11

AAx87319
ID AAx87319 standard; cDNA: 754 BP.

AX AX87319;

XX 27-SEP-1999 (first entry)

XX Human sensitive to apoptosis (SAG) gene mutant MM3.

XX SAG gene: sensitive to apoptosis; human; cancer; tumour;

KW neurodegenerative disease; muscular dystrophy; wound healing;

XX vulnerable; therapy; mutant; ds.

OS Homo sapiens.

XX Synthetic.

XX Key Location/Qualifiers

FT CDS 1..342

FT mutation /tag= a

FT mutation /note= "C50S mutation"

FT mutation /tag= c

FT mutation /note= "C53S mutation"

FT mutation /note= "C53S mutation"

FT mutation /note= "C53S mutation"

FT mutation /note= "C53S mutation"

FT mutation /note= "C53S mutation"

FT mutation /note= "C53S mutation"

FT mutation /note= "C53S mutation"

FT mutation /note= "C53S mutation"

FT mutation /note= "C53S mutation"

PT promoting cell growth and protecting cells against apoptosis
XX
PS Claim 15; Page 62-63; 84pp; English.
XX
CC This is the nucleotide sequence of human sensitive to apoptosis
CC (SAG) mutant gene MM3, which codes for a SAG protein (see AA106497)
CC in which the Cys residues at positions 50 and 53 of the native
CC protein (see AA106492) are replaced by Ser residues owing to
CC mutations of codons 50 and 53 from TGC to AGC obtained by
CC site-directed mutagenesis of SAG cDNA. These residues are at a haem
CC binding site of SAG. Single and double SAG mutants (see AAx87317-31)
CC were made in order to determine the role of each cysteine residue
CC of SAG in haem binding and SAG oligomerization. MM3 showed greatly
CC reduced haem binding but oligomerization was unaffected. SAG is a
CC novel zinc finger protein that promotes cell growth, protects cells
CC from apoptosis, scavenges oxygen radicals and can be used for the
CC reversion of a tumour phenotype. SAG genes, and mutant SAG genes,
CC can be used to protect cells from apoptosis induced by redox
CC reagents. They can also be used for the recombinant production of
CC SAG proteins, which are molecular targets in the development of
CC drugs against neurodegenerative disorders, cancers and muscle
CC dystrophy, and promoting wound healing.
XX
SQ Sequence 754 BP; 207 A; 155 C; 201 G; 191 T; 0 other;

Query Match 99.6%; Score 750.8; DB 20; Length 754;
Best Local Similarity 99.7%; Pred. No. 2.4e-230;
Matches 752; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCCGAGTGGAGAGAGAGAGAAACCTGGCCCTGCTCTACCTCCGGAGCTCA 60
| | | | |
DB 1 atggccgagctggagagagagagaaacctggccctgctctacccctccggagctca 60
QY 61 GGCTCCAAAGTGGGAGGCGCAAGATGTTCTCCCTCAAGAGTGGAGCGGCTGCAATG 120
| | | | |
DB 61 ggctccaaagtgggagggagcaagaatgttctccctcaagaagtggagcggtggcag 120
QY 121 TGGAGCTGGGAGCTGGAGTGGCGATCGTGGCCATCGAGGATGATGATGATGATGATG 180
| | | | |
DB 121 tggagctgggagctggagtggtgagcagcagccatcagaaggtgagtgatgagtgatg 180
QY 181 TGTCTTAGATGTCAGCTGAGTGAACCAAGAGAGACTGTTGTGCTGGGGAGATGT 240
| | | | |
DB 181 tgtcttagatgtcagctgagtgagtcagtcagtcagtcagtcagtcagtcagtcagtc 240
QY 241 AATCATTCCTTCCACAAAGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
| | | | |
DB 241 aatcatctctccacaaagctgctgagtcagtcagtcagtcagtcagtcagtcagtcag 300
QY 301 CTCTCCACAGAGAGCTGGGTGTCCTCAAGAGATGCGCAATGAGAGTGTAGAGAGCTTC 360
| | | | |
DB 301 ctctccacagagagctgggtgtctcctcaagagatgcgcaatgagagtgtagagagcttc 360
QY 361 TTAGCGCAGTTGTTACAGACCCCTGGTGATCTTGTATATCAGTGCCTTCAAGAGCTTCA 420
| | | | |
DB 361 tttagcgagttgttacagacccctggtgatcttgtatattcagtcagtcagtcagtcag 420
QY 421 ACACTACAGGGAGTGAATCTTCAATAGAGCGGATGATGATGATGATGATGATGATGAT 480
| | | | |
DB 421 acctacagggagtgaaatcttcaatagagcggtgagtcagtcagtcagtcagtcagtcag 480
QY 481 CAAAGCCTTGTTAGCATTTGTGAGTTTATCTTACAGAAATTCCTGTGATTAAGAAGAT 540
| | | | |
DB 481 caaagccttgtagcatattgtgagtttattcttcaagaattctctgtgattaaagaagt 540
QY 541 AATTATTAAGGTGCTCTCTACCTCTGTGGTGTGTGTGCGCACACAGCTTAGAAG 600
| | | | |
DB 541 aattatlaaaggtygtccctcctaccctcgtgtgtgtcgcgcacacagcttagaag 600
QY 601 TGGTATTAAGAAAGAAAGAGCTCCAAATTAATCACCCTTAAATTTACCATTTCTATAC 660
| | | | |
DB 601 tgcataaaaaaagaaagagctccaattgatacccttaatttaccattctctatc 660

OY 661 AACAGCAGTGGAGACGTTTCGACGTTTGCATTGATGATCACTTAATAAAA 720
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 661 aacagcagtggaagcagcttcgagacttctcgactccttaagctgacgttaaaaa 720
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 OY 721 GAATGTTACAGTACAAATAAAGTCAGTTTAAA 754
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 721 gaatgtacagtaacaataaagtcagtttaa 754
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 RESULT 12
 AAX87322 standard; cDNA; 754 BP.
 XX AAX87322;
 AC AAX87322;
 XX 27-SEP-1999 (first entry)
 DE Human sensitive to apoptosis (SAG) gene mutant MM6.
 XX
 XX SAG gene; sensitive to apoptosis; human; cancer; tumour;
 KW neurodegenerative disease; muscular dystrophy; wound healing;
 KM vulnerability; therapy; mutant; ds.
 OS Homo sapiens.
 OS Synthetic.
 XX
 PH Key Location/Qualifiers
 FT CDS 1..342
 FT /*tag= a
 FT replace(244..246,CAT)
 FT /*tag= b
 FT /note= "H82K mutation"
 FT XX
 PN WO932514-A2.
 XX
 PD 01-JUL-1999.
 XX
 PD 15-DEC-1998; 98WO-US26705.
 PF
 XX 11-SEP-1998; 98US-0099840.
 PR 19-DEC-1997; 97US-0068179.
 XX
 PA (WARN) WARNER LAMBERT CO.
 XX
 PI Sun Y;
 DR WPI; 1999-430152/36.
 DR P-PSDB; AAY06500.
 XX
 PT SAG: Sensitive to Apoptosis Gene and related proteins, useful for
 PT promoting cell growth and protecting cells against apoptosis
 XX
 PS Claim 15; Page 67-68; 84pp; English.
 XX
 XX This is the nucleotide sequence of human sensitive to apoptosis
 CC (SAG) mutant gene MM6, which codes for a SAG protein (see AAY06500)
 CC in which the His residue at position 82 of the native protein (see
 CC AAY06452) is replaced by a Lys residue owing to a mutation of codon
 CC 82 from CAT to AAA obtained by site-directed mutagenesis of SAG
 CC cDNA. This residue is in zinc-finger 1 of SAG. Single and
 CC double SAG mutants (see AAX87317-31) were made in order to determine
 CC the role of each cysteine residue of SAG in haem binding and SAG
 CC oligomerization. These properties were unaffected by the MM6
 CC mutation. SAG is a novel zinc finger protein that promotes
 CC cell growth, protects cells from apoptosis, scavenges oxygen
 CC radicals and can be used for the reversion of a tumour phenotype.
 CC SAG genes, and mutant SAG genes, can be used to protect cells from
 CC apoptosis induced by redox reagents. They can also be used for the
 CC recombinant production of SAG proteins, which are molecular targets
 CC in the development of drugs against neurodegenerative disorders,
 CC cancers and muscle dystrophy, and promoting wound healing.
 XX

SQ Sequence 754 BP; 207 A; 154 C; 201 G; 192 T; 0 other;
 Query Match 99.6%; Score 750.8; DB 20; Length 754;
 Best Local Similarity 99.7%; Pred. No. 2.4e-230;
 Matches 752; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 OY 1 ATGGCGACGTGGAAGACGAGGAAACCTCGCGCCTGCTCCTACCTCGGAGCTCA 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 1 atggcgcagctggaagacggaagaaacccgcccgcctcccaatccgcggagctca 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 OY 61 GGCTCCAAAGTCGGGAGGCGACAAAGATGTTCTCCCTCAAGAACTGGAACGCGGTGCAT 120
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 61 ggctccaaagtcgggagcggaagaagtgcttcctcccaagaagtggaacgcggtgcctatg 120
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 OY 121 TGGAGCTGGGAGAGCTGGAGTGCATATGCTGCCATCTGAGAGGTCCAGGTATGATGTC 180
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 121 tggagctgggagagctggagtgcatatgctgccatctgagaggtccaggtatgattgctc 180
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 OY 181 TGTCTTAGATGTCAAGCTGAATAAACAAGAGAGACTGTGTGTGGTGGGAGAAATGT 240
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 181 tgtcttagatgtcaagctgaataaacaagagagactgtgtgtgggtgggagaaatgtgt 240
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 OY 241 AATCATTCCTTCACAACTGCTGCATGTCCTGTGGGTAAACAGAAATGCGTGCCT 300
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 241 aatcatctcttcacaaactgctgcattgctcctgtgggtaaacagaaatgctgcctct 300
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 OY 301 CTCTGCACAGAGACTGGGTGTCCAAAGATCGCAATGAGATGGTTAAGGCTTC 360
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 301 ctctgcacagagactgggtgtccaaagatcgcaatgagatgggttaagagcttc 360
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 OY 361 TTAGCGCAGTGTTCAGAGCCCTGTGTGATCTTGAATCAGATGCCCTCAAGAGCTAGA 420
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 361 tttagcgagttgtcagagccctgtgtgatcttgaatcagatgccctcaagagctaga 420
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 OY 421 ACAGTACAGGGAGTGAATTCCTCAATATAGAGCCGATGAGTGTGCTTTGACTCAT 480
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 421 acactacagggagtgatctctcaaatatagagccgatgagtgctttgactcat 480
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 OY 481 CAAAGCCTGTTAGATTTGACGTTTCACTTATCTCAGAAATCTCTGATTAAGAGAT 540
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 481 caaagcctgttagatttgcagtttcaacttattctcgaatctctgtatcaagaagat 540
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 OY 541 AATTATTAAAGTGCTCTCTCTACCTGCTGTGTGTGTGCGCACACAGCTTAGAAG 600
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 541 aattattaaagtgtctctctctacacctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 600
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 OY 601 TGCTATTAATAAAGAAAGAGCTCAATTAATGAATACCTTATTAATTTACCAATTTCTATAC 660
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 601 tgctataaataaagaaagagctccaattgaatcaccattataatccattctatag 660
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 OY 661 AACAGCAGTGGAGACGTTTCGAGACTTTTCGATGCTTAAAGTGAACACTTAATAAAA 720
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 661 aacagcagtggaagcagcttcgagacttctcgactccttaagctgacgttataaaaa 720
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 OY 721 GAATGTTACAGTACAAATAAAGTCAGTTTAAA 754
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 721 gaatgtacagtaacaataaagtcagtttaa 754
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 RESULT 13
 AAX87323 standard; cDNA; 754 BP.
 XX AAX87323;
 AC AAX87323;
 XX 27-SEP-1999 (first entry)
 DE Human sensitive to apoptosis (SAG) gene mutant MM7.
 XX
 XX SAG gene; sensitive to apoptosis; human; cancer; tumour;
 KW neurodegenerative disease; muscular dystrophy; wound healing;
 KM vulnerability; therapy; mutant; ds.
 KW

XX OS Homo sapiens.
 XX OS Synthetic.
 XX FH Key
 FT CDS Location/Qualifiers
 FT 1..342
 FT /tag= a
 FT replace(253..255,CAC)
 FT mutation /tag= b
 FT /note= "H85K mutation"
 XX

W09932514-A2.

01-JUL-1999.

15-DEC-1998; 98WO-US26705.

PR 11-SEP-1998; 98US-0099840.
 PR 19-DEC-1997; 97US-0068179.

(WARN) WARNER LAMBERT CO.

Sun Y;

WPI: 1999-430152/36.

P-PSDB; AAY06501.

SAG: Sensitive to Apoptosis Gene and related proteins, useful for promoting cell growth and protecting cells against apoptosis
 Claim 15; Page 69-70; 84pp: English.

This is the nucleotide sequence of human sensitive to apoptosis (SAG) mutant gene MM7, which codes for a SAG protein (see AAY06501) in which the His residue at position 85 of the native protein (see AAY06492) is replaced by a Lys residue owing to a mutation of codon 85 from CAC to AAG obtained by site-directed mutagenesis of SAG cDNA. This residue is in zinc-finger finger 2 of SAG. Single and double SAG mutants (see AAX87317-31) were made in order to determine the role of each cysteine residue of SAG in haem binding and SAG oligomerization. These properties were unaffected by the MM7 mutation. SAG is a novel zinc finger protein that promotes cell growth, protects cells from apoptosis, scavenges oxygen radicals and can be used for the reversion of a tumour phenotype. SAG genes, and mutant SAG genes, can be used to protect cells from apoptosis induced by redox reagents. They can also be used for the recombinant production of SAG proteins, which are molecular targets in the development of drugs against neurodegenerative disorders, cancers and muscle dystrophy, and promoting wound healing.

Sequence 754 BP; 206 A; 153 C; 202 G; 193 T; 0 other;

Query Match 99.6%; Score 750.8; DB 20; Length 754;
 Best Local Similarity 99.7%; Pred. No. 2.4e-230;
 Matches 752; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCCGACGTGGAAGAAGGAGAGAAACCTGGCCCTCTCACTCCGGAGCTCA 60
 DB 1 atggccgacgtggaaagcggagaaacctgcccctcctcaactccgagagctca 60
 QY 61 GGCTCCAAAGTCGGAGCGACGAAGATGTTCTCCCTCAAGAAGTGAAGCGGGGCGCAG 120
 DB 61 ggcctccaaagtcggagcgacgaaagatgttctccctcaagaagtggaacgaggcgccag 120
 QY 121 TGGAGCTGGGACGTGAGTGCATACGTGCGCCATCTGCAGGCTCCAGGTGATGATGCC 180
 DB 121 tggagctgggacgtgagtgcatatcgatgcacatctgcagggctccagtgatgatgcc 180
 QY 181 TGGCTTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
 DB 181 tggcttagatgatatgatatgatatgatatgatatgatatgatatgatatgatatgatatg 240

QY 241 AATCATCTCTTCCACAACTGCTGCATGCTCCCTGTGGGTGAAGAAACAATCCGTGCCCT 300
 DB 241 aatcatctcttccacaaactgctgcatagtctccctgtgggtgaagaaacaatccgtgccct 300
 QY 301 CTCTGCCAGCAGGACCTGGGTGTCACAAAGATGGGCAATGACAGTGTGTAAGAGCTTC 360
 DB 301 ctctgccagcagacctgggtgtctacaaagatgggcaatgacagtgtgtgtaagagcttc 360
 QY 361 CTCTGCCAGCAGGACCTGGGTGTCACAAAGATGGGCAATGACAGTGTGTAAGAGCTTC 420
 DB 361 ctctgccagcagacctgggtgtctacaaagatgggcaatgacagtgtgtgtaagagcttc 420
 QY 421 TTAGCGGAGTGTGTGACAGCCCTGTGGATCTTGTATCCAGTCCCTCAAGAGCTTAGA 480
 DB 421 ttagcgagtggtgtgacagccctgtggatcttgtatccagtccctcaagagcttaga 480
 QY 481 CAAAGCCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
 DB 481 caaagccttgtgatgatgatgatgatgatgatgatgatgatgatgatgatgatgatgat 540
 QY 541 AATTTTAAAGTGTGCTCTTCTACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
 DB 541 aattttaaagtgtgctcttctacctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 600
 QY 601 TGCTATTAAGAAAGGAAAGAGCTCCAAATGAATGAAATGAAATGAAATGAAATGAAAT 660
 DB 601 tgctatataaagaaagagctccaaatgaatgaatgaatgaatgaatgaatgaatgaatga 660
 QY 661 AACAGCAGTGAAGAGAGCTTTCGAGACTTTTTCGATGATGATGATGATGATGATGAT 720
 DB 661 aacagcagtgaagagagcttctcgagacttttctcgatgatgatgatgatgatgatgatga 720
 QY 721 GAATGTTACAGTACCAATTAAGTGCAGTTTAA 754
 DB 721 gaatgttacagtaccaatataagtgcagtttaa 754

RESULT 14

AAX87329
 ID AAX87329 standard; cDNA; 754 BP.

AC AAX87329;

DT 27-SEP-1999 (first entry)

DE Human sensitive to apoptosis (SAG) gene mutant MM13.

KW SAG gene; sensitive to apoptosis; human; cancer; tumour;

KW neurodegenerative disease; muscular dystrophy; wound healing;

KW vulnery; therapy; mutant; ds.

OS Homo sapiens.

OS Synthetic.

FT Key

FT CDS

FT mutation Location/Qualifiers
 /tag= a
 /replace(181,T)
 /tag= b
 /note= "C61S mutation"
 /replace(190,T)
 /tag= c
 /note= "C64S mutation"

PN W09932514-A2.

PD 01-JUL-1999.

PF 15-DEC-1998; 98WO-US26705.

PR 11-SEP-1998; 98US-0099840.

PR 19-DEC-1997; 97US-0068179.

PA (WARN) WARNER LAMBERT CO.

PI Sun Y;

DR WPI; 1999-430152/36.

DR P-PSDB; AAY06507.

PT SAG: Sensitive to Apoptosis Gene and related proteins, useful for promoting cell growth and protecting cells against apoptosis

PS Claim 15; Page 79-80; 84pp; English.

This is the nucleotide sequence of human sensitive to apoptosis (SAG) mutant gene MM13, which codes for a SAG protein (see AY06507) in which the Cys residues at positions 61 and 64 of the native protein (see AY06492) are replaced by Ser residues owing to mutations of codons 61 and 64 through site-directed mutagenesis of SAG cDNA. These residues are in zinc ring finger 1 of SAG. Single double SAG mutants (see AA87317-31) were made in order to determine the role of each cysteine residue of SAG in haem binding and SAG oligomerization. MM13 showed greatly reduced haem binding (and a decreased brown colour) but oligomerization was unaffected. SAG is a novel zinc finger protein that promotes cell growth, protects cells from apoptosis, scavenges oxygen radicals and can be used for the reversion of a tumour phenotype. SAG genes, and mutant SAG genes, can be used to protect cells from apoptosis induced by redox reagents. They can also be used for the recombinant production of SAG proteins, which are molecular targets in the development of drugs against neurodegenerative disorders, cancers and muscle dystrophy, and promoting wound healing.

SQ Sequence 754 BP; 207 A; 155 C; 201 G; 191 T; 0 other;

SQ Sequence 754 BP; 207 A; 155 C; 201 G; 191 T; 0 other;

SQ Sequence 754 BP; 207 A; 155 C; 201 G; 191 T; 0 other;

Query Match	99.68;	Score 750.8;	DB 20;	Length 754;
-------------	--------	--------------	--------	-------------

```

Matches 752; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY	1	ATGCGCAGCTGGAAAGCAGAGAGAAACCTGCGCCCTGCTCTCTACTCCGGAGCTTA	60
Db	1	atgcccagcttggaaagcaggagagaaacctgcgccttgccttcaactccggagctca	60
QY	61	GGGCCCAAGTCGGGAGCGCGCAAGAGTGTCTCCCTCAAGAAAGTGAAGCGGGGGCAAG	120
Db	61	ggcccacaagctcggagagcgacgaagaagcttccctcaagaagtggaagcggtgtgcacag	120
QY	121	TGGAGCTGGAGCACTGGAAGTCCGATACGTGCGCCATCTGCAGGGTCCAGGTGATGATGCC	180
Db	121	tggagctcggagagctggagatgctgaagctgcgcacatctcgaagtggtccagatgatgatgcc	180
QY	181	TGTCCTTAGATGTCAGCGCTGAACCAACAAGAGAGACTGTGTGTGGTCTGGGAGAAATGT	240
Db	181	agctcttgaagtaagctggaacaacaagaagagactgtgtgtgtcttggggagaagt	240
QY	241	AATCATTCCTTCCACAACACTGCTGATGTCCCTGTGGGTGAACAACAATGCTGCTCCCT	300
Db	241	aatcatctcttccacaactctgtcagctccctgttgggtgaacaacaatctgtccct	300
QY	301	CTCTGCCAGAGAGACTGGGGTGCACAAAGAAATGGCAATGAGAGGTGTTAAGAGCTTC	360
Db	301	ctctgcgcagaagagactggggtgcacaaagaaatggcaatgagaggtgttlaaagaagcttc	360
QY	361	TTAGCGCAGTGTTCACAGAGCCGTGTGGATCTGTATCCAGTGCCTCCAAAGGCTAGA	420
Db	361	ttagcgcagttgttcacagagccgtgtggatctgtatccagtgcctcaagaagctaga	420
QY	421	ACACTACAGGGGATGAATTTCTCAAAATAGAGCCGATGATCTGTGCTTTTGACATCAT	480
Db	421	aacactacagggataagtaattctcaaatagagccgatgatctgtgtcttvtgagcat	480
QY	481	CAAAAGCCTTGATAGCAATTTGTCAGATTTATCTTCGAAGAAATCTCGTATTAAGAAGAT	540
Db	481	caaagccttggatagcaatttgtcagtttatcttcgaagaattctcgtatlaagaagat	540

QY	541	AAATATTAAAGGCGGCGCTTCTCACTCGTGTGGTGTGTGCGGGCACACAGTTAAAG	600
QY	541	aaattactaaagaaggagccttcctccaccctcggggtcggtgcgcacacagctlaaag	600
Db	541	aaattactaaagaaggagccttcctccaccctcggggtcggtgcgcacacagctlaaag	600
QY	601	TGCTATTAATAAAGGAAGAAGAGCTCCAAATTAATCACTTAAATTAATACCATTTCTTAC	660
Db	601	tgctataaaaaaggaaagaagagctcccaattgaatcaacttaattatcccaattctatac	660
QY	661	AACAGCAGTGGGAAGCAGGTTTCGAGACTTTTTCGATGCTTAATGGTGCATGATTAAAAA	720
QY	661	aacagcgagtcggaagcaggttcgcgagacttttcgatcgtctatgltgatcagtaaaaaa	720
Db	661	aacagcgagtcggaagcaggttcgcgagacttttcgatcgtctatgltgatcagtaaaaaa	720
QY	721	GAATGTTACAGTAAACAATTAATGAAGCGAGTTTAAA	754
QY	721	gaatglttacagtlaacaataaaggagcggtctaaa	754

RESULT 15

ID	AA	CDNA	BP
AA	AX87330	standard	754

AC AAX87330;

DT 27-SEP-1999 (first entry)

Human sensitive to apoptosis (SAG) gene mutant MML4.

KW SAG gene; sensitive to apoptosis; human; cancer; tumour;

KW vulnerable; therapy; mutant; ds.

Homo sapiens.

XX XX

FT	CDS	1..342
----	-----	--------

```
FT      mutation      replace(295,c)
```

```

/note= "C99S mutation"

```

```
FT      /*tag= c 01000 2014:0
```

XX

XX
XX

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

XX
XX

PR 19-DEC-1997; 97US-0068179.

PA (WARN) WARNER LAMBERT CO.

PI Sun Y;

DR WPI; 1999-430152/36.

[illegible]

PT promoting cell growth and protecting c

PS Claim 15; Page 81; 84pp; English.

CC This is the nucleotide sequence of human
CC (316) mutant gene wv14 which codes for

CC in which the Cys residues at positions

CC mutations of codons 99 and 102 through
CAC CTAA These residues are in the

CC double SAG mutants (see AAX87317-31) were made in order to determine

Search completed: July 25, 2002, 20:25:51
Job time: 10974 sec

Query Match	99.6%	Score 750.8;	DB 20;	Length 754;
Best Local Similarity	99.7%;	Pred. No. 2.4e-230;		
Matches 752;	Conservative	0;	Mismatches 2;	Indels 0;
			Gaps	0

QY	1	ATGCGCAGCTGGAAACACGAGAGGAAACCTGGCCCTGGCCCTGCACCTCCGGAGCTCA	60
Db	1	atggcgcagcttgaaagacggaagaaacctgcgccttgcttccatctcggggagctcca	60
QY	61	GGCTCCAAAGTCGGGAGGCGCAAGAGTCTCTCCCTCAAGAAATGGAACCGGGTGCATG	120
Db	61	ggctccaaagtcgggagggcgacaagaatgcttctccctcaagaagtggaaagcggtyggcaty	120
QY	121	TGGAGCTGGGACGTGGAGTGCATACGTGCGCCATCTGCAGGGTCCAGGTGATGGATGCC	180
Db	121	tggagctgggacgtggagtgcatatcgatgcgcacatctgcaagtgatccaagtgatgatgcc	180
QY	181	TGCTTTGAGTGTCAAGCTGAAAAAACAAGAGAGACTGTGTGTGTGTGTGGGAGAAATGT	240
Db	181	tgctttgagtgtcaagcttgaaaaacaagaagagactgtgtgtgtgtgtgtgtgggagaatgt	240
QY	241	AATCATTTCTTCCACAACCTGCTGCATGTCCCTGTGGGTGAACAGAACATGCTGCCCT	300
Db	241	aatcatcttctccaacactctgcatgtccctgtgggtgaacagaaacatgcagccct	300
QY	301	CTCTGCCACAGACACTGGGTGTCCAAAGAAATGGCAAAAGAGAGGTATAGAGGCTTC	360
Db	301	ctctgccacagacactgggtgtccaaagaaatggcaaaagagagtggtatagaagcttc	360
QY	361	TTAGGGCAGTTGTTCCAGAGCCCTGGTGAATCTGTATCCAGTGCCTTACMAAGGCTAGA	420
Db	361	ttagggcagttgttccagagccctggtgaaatctgtatccagtgccttacmaaggctaga	420
QY	421	ACACTACAGGGGATGAATTTCTTCAATATAGAGCCGATGATCTGTGTCTTTTGACTACT	480
Db	421	acactacaggggatgaatttcttcaaatatagagccgatgatctgtgtcttttgactact	480
QY	481	CAAAAGCTTTGGTTAGCATTTTGTCAAGTTTAACTTCACAAATTCCTGTGATTTAAGAAAT	540
Db	481	caaaagctttggttagcatttgcagtttacttccagaaatcttcgtgattagaagat	540
QY	541	AATTTATTTAAAGGTGTGCTCTTCAACCTGTGTGTGTGTGTGCGGCAACAGCTTAGAG	6000
Db	541	aatttatTTAAAGGTGTGCTCTTCAACCTGTGTGTGTGTGTGCGGCAACAGCTTAGAG	6000
QY	601	TGCTTTAAAAAAGAAAGAGCTCCAAATTTGAATCACTTATTAATTTAACCATTTCTATAC	6600
Db	601	tgccttaaaaaagaaagagctccaaatTTGAATCACTTATTAATTTAACCATTTctatac	6600
QY	661	AACAGGAGGTGAGACACTTTTCGAGACTTTTTCGATGTTATGGTTATCAGTTAAAAAA	7200
Db	661	aacaggaggtgagacagcttttcgagacttttTCGATGTTATGGTTATCAGTTaaaaa	7200
QY	721	GAAATGTTACAGTAACAATAAATTAAGGTCAGTTAAA 754	4
Db	721	gaatgttacagtaacaaaaaagtgagtttaa 754	

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 25, 2002, 14:08:41 ; Search time 3583.2 Seconds

(Without alignments)
6657.806 Million cell updates/sec

Title: US-09-509-779-1

Perfect score: 1140

Sequence: 1 GTCCTGCGCGCGCGCCATG.....ACTATTCAATCAATTAAG 1140

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :
1: GenEmbl:
2: gb_ba:
3: gb_hcg:
4: gb_in:
5: gb_om:
6: gb_ov:
7: gb_pat:
8: gb_ph:
9: gb_pl:
10: gb_pr:
11: gb_ro:
12: gb_scs:
13: gb_sy:
14: gb_un:
15: gb_vl:
16: em_ba:
17: em_fun:
18: em_hum:
19: em_in:
20: em_mu:
21: em_om:
22: em_or:
23: em_ov:
24: em_pat:
25: em_ph:
26: em_pl:
27: em_ro:
28: em_sts:
29: em_un:
30: em_vl:
31: em_htg_hum:
32: em_htg_inv:
33: em_htg_other:
34: em_htg_inv:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
------------	-------	-------------	--------	-------	-------------

1	1140	100.0	1140	10	AF092877	AF092877 Mus muscu
2	1064.6	93.4	1125	10	BC011127	BC011127 Mus muscu
3	364.4	32.0	816	9	BC008627	BC008627 Homo sapi
4	362.8	31.8	836	9	AF164679	AF164679 Homo sapi
5	355.4	31.2	822	9	BC005966	BC005966 Homo sapi
6	348.4	30.6	163521	2	AC108040	AC108040 Homo sapi
7	348.4	30.6	163757	2	AC092418	AC092418 Homo sapi
8	334	29.3	754	9	AF092878	AF092878 Homo sapi
9	306.6	26.9	53155	2	AC104983	AC104983 Homo sapi
10	303.2	26.8	342	6	AF142060	AF142060 Homo sapi
11	303.6	26.6	342	6	AX282592	AX282592 Sequence
12	256.4	22.5	170087	2	AC098496	AC098496 Rattus no
13	236.6	20.8	169675	9	AC060766	AC060766 Homo sapi
14	183.4	16.1	160012	2	AC106176	AC106176 Rattus no
15	164.2	14.4	439	9	HSB333208	HSB333208 Homo sapi
16	164.2	14.4	179877	2	AC092978	AC092978 Homo sapi
17	164.2	14.4	191221	2	AC067952	AC067952 Homo sapi
18	162.6	14.3	674	9	AF312226	AF312226 Homo sapi
19	158.4	13.9	125041	2	AC068693	AC068693 Homo sapi
20	158	13.9	191221	2	AC067952	AC067952 Homo sapi
21	96	8.4	125041	2	AC068693	AC068693 Homo sapi
22	90.6	7.9	327	9	AF142059	AF142059 Homo sapi
23	90.6	7.9	497	9	HOMO60A05	HOMO60A05 Homo sapi
24	90.6	7.9	508	9	AF140598	AF140598 Homo sapi
25	90.6	7.9	544	9	BC001466	BC001466 Homo sapi
26	90.6	7.9	554	9	BC017370	BC017370 Homo sapi
27	90.6	7.9	5347	6	AX281690	AX281690 Sequence
28	87.8	7.7	88787	2	AC020266	AC020266 Drosophila
29	87.8	7.7	170869	3	AC011696	AC011696 Drosophila
30	87.8	7.7	171831	3	AC007473	AC007473 Drosophila
31	87.8	7.7	278196	3	AE003825	AE003825 Drosophila
32	84	7.4	553	3	AY070810	AY070810 Drosophila
33	84	7.4	750	3	AF218290	AF218290 Drosophila
34	84	7.4	3185	3	AY061302	AY061302 Drosophila
35	84	7.4	95294	2	AC019742	AC019742 Drosophila
36	84	7.4	167962	3	AC010010	AC010010 Drosophila
37	84	7.4	310120	3	AE003468	AE003468 Drosophila
38	83.8	7.4	3484	9	HSTEST	HSTEST H. sapiens m
39	83.4	7.3	3208	6	AX212267	AX212267 Sequence
40	79.4	7.0	504	10	AF140599	AF140599 Mus muscu
41	76.6	6.7	389	6	AX192664	AX192664 Sequence
42	75.4	6.6	388	6	AX192698	AX192698 Sequence
43	71.2	6.2	361	3	AB077287	AB077287 Caenorhab
44	70.6	6.2	324	8	AF179228	AF179228 Schizosac
45	70.6	6.2	35277	3	CBRG19102	CBRG19102 Caenorhab

ALIGNMENTS

RESULT 1
LOCUS AF092877 1140 bp mRNA linear ROD 24-JUL-2001
DEFINITION Mus musculus zinc RING finger protein SAG mRNA, complete cds.
ACCESSION AF092877
VERSION AF092877.1 GI:4588031
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Scuriognathi; Muridae; Murinae; Mus.

REFERENCE
1 (bases 1 to 1140)
Swaroop,M., Bian,J., Aviram,M., Duan,H., Bisgater,C.L., Loo,J.A.
and Sun,Y.
Expression, purification, and biochemical characterization of SAG,
a RING finger redox-sensitive protein
Free Radical Biol. Med. 27, 193-202 (1999)

TITLE
JOURNAL
AUTHORS
REFERENCE
2 (bases 1 to 1140)
Sun,Y.
Alterations of SAG mRNA in human cancer cell lines: requirement for
the RING finger domain for apoptosis protection
Carcinogenesis. 20 (10), 1899-1903 (1999)

JOURNAL
MEDLINE
99435944

QY	241	GGTCGGGGAGAGGTAAACCAATTCCTCCACAACTGCATGTCGCCGTGGGGAAACA	300
Db	241	GGTCGGGGAGAGGTAAACCAATTCCTCCACAACTGCATGTCGCCGTGGGGAAACA	300
QY	301	GAACAATCGTCAGCCCTCTGTCCACAGAGACTGGGTAGTCCMAAGAAATCGGCAATGAGA	360
Db	301	GAACAATCGTCAGCCCTCTGTCCACAGAGACTGGGTAGTCCMAAGAAATCGGCAATGAGA	360
QY	361	GGTGGCCAGGCGCTCTGGTGTGGTTGCTGACCTGGACAAAGACTTAAACCTGCGAGG	420
Db	361	GGTGGCCAGGCGCTCTGGTGTGGTTGCTGACCTGGACAAAGACTTAAACCTGCGAGG	420
QY	421	GATTCATCCTTGAGAGAGAGAGATGCTGTGCGCCTTGTGAACATCACCACCAAGCGTTGCTT	480
Db	421	GATTCATCCTTGAGAGAGAGAGATGCTGTGCGCCTTGTGAACATCACCACCAAGCGTTGCTT	480
QY	481	TATTAATATTGTCTGTTTAGTTTTGGGAAATTCCTACATTTAAGATTAATTTGTTAAAAAT	540
Db	481	TATTAATATTGTCTGTTTAGTTTTGGGAAATTCCTACATTTAAGATTAATTTGTTAAAAAT	540
QY	541	GGCCCTTCCCTACCTCGGTGGTGTGGTGTGATAGAAATGATAGAGAGCGAGAACCA	600
Db	541	GGCCCTTCCCTACCTCGGTGGTGTGGTGTGATAGAAATGATAGAGAGCGAGAACCA	600
QY	601	GAATAATGATCTTTGTTTATCTGTATCCCAAGCACTGGAACTGTGTTCCAGAGAATCTT	660
Db	601	GAATAATGATCTTTGTTTATCTGTATCCCAAGCACTGGAACTGTGTTCCAGAGAATCTT	660
QY	661	GTTTGTGTTTATGCTGTGAGGGTTAAAAATAGATTAAGAAATGTTACGTAAACAATATA	720
Db	661	GTTTGTGTTTATGCTGTGAGGGTTAAAAATAGATTAAGAAATGTTACGTAAACAATATA	720
QY	721	ATGCATTAAGAAAGCGAGCTCCTCTAATCTTTTGTGTGGAGAGAGGCGAAGCGAGCG	780
Db	721	ATGCATTAAGAAAGCGAGCTCCTCTAATCTTTTGTGTGGAGAGAGGCGAAGCGAGCG	780
QY	781	CACCCGTGTCCTTCAATTTGCTGTGATAGAGATTTTAACCTGCATCAGTGAAGAGCGG	840
Db	781	CACCCGTGTCCTTCAATTTGCTGTGATAGAGATTTTAACCTGCATCAGTGAAGAGCGG	840
QY	841	TAACTGTGGGTAACCTGTAATGCGGTACTGTGCGGGTAAAGGGCTTGTCTCTGAC	900
Db	841	TAACTGTGGGTAACCTGTAATGCGGTACTGTGCGGGTAAAGGGCTTGTCTCTGAC	900
QY	901	TTCCTCATCTTTGACTTGGCCAGAGCGCTGGATTGTCTCAACCACTTAGTCTTAAAGAAC	960
Db	901	TTCCTCATCTTTGACTTGGCCAGAGCGCTGGATTGTCTCAACCACTTAGTCTTAAAGAAC	960
QY	961	TGTTTCTGTTTGTGGCCAGAGGTGTATTGTATGTTTATGTCACAAAATATTAGTGAAGA	1020
Db	961	TGTTTCTGTTTGTGGCCAGAGGTGTATTGTATGTTTATGTCACAAAATATTAGTGAAGA	1020
QY	1021	ATGGCTATCTAGTATTAACAGAGAACTGATTAATGCAATGTTTAAATAAATTTGCTT	1080
Db	1021	ATGGCTATCTAGTATTAACAGAGAACTGATTAATGCAATGTTTAAATAAATTTGCTT	1080
QY	1081	TGAGTTATTAAAGTTTGAATATATACTCTTAAATCATTTAAACTAATTCATCAATTAATG	1140
Db	1081	TGAGTTATTAAAGTTTGAATATATACTCTTAAATCATTTAAACTAATTCATCAATTAATG	1140
RESULT 2			
LOCUS	BC011127	1125 bp	mRNA linear ROD 30-JUL-2001
DEFINITION	Mus musculus, similar to ring finger protein 7, clone MGC:19076		
ACCESSION	BC011127		
VERSION	BC011127.1	GI:15029807	
KEYWORDS	MGC.		
SOURCE	house mouse.		
ORGANISM	Mus musculus.		
	Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;		

REFERENCE

1 (bases 1 to 1125)
AUTHORS
TITLE
JOURNAL

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Strausberg, R.
Direct Submission
Submitted (25-JUL-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK

NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
DNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: villalona@bcm.tmc.edu
Villalón, D.K., Luna, R.A., Hale, S.M., Huliyk, S., Lu, X., Garcia,
A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,
Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/BLIN at: <http://image.llnl.gov>
Series: IRAC Plate: 24 Row: k Column: 5
This clone was selected for full length sequencing because it
passed the following selection criteria: GenomScan gene
prediction, similarity but not identity to protein.
Location/Qualifiers

FEATURES

source

CDS

1..1125
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="MGC:19076 IMAGE:4194107"
/tissue="Liver, normal, 5 month old male mouse."
/clone_lib="NCI_CGAP_L19"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
12..353
/codon_start=1
/product="Similar to ring finger protein 7"
/protein_id="AAH1127.1"
/db_xref="GI:15029808"
/translation="MADVEDGEPCVLSHSGSAGSKSGDKMFSLKKNVAVMMSWD
VECDPFCACRVOVMDACLRQAKNKEDECVVWGECHSHFHNCSMLWVQNNRCPIC
QQDWVYORIRK"
BASE COUNT 304 a 219 c 286 g 316 t
ORIGIN

Query Match 93.4% Score 1064.6 DB 10 Length 1125;
Best Local Similarity 99.18; Pred. No. 1.9e-259;
Matches 1113; Conservative 0; Mismatches 4; Indels 6; Gaps 4;
QY 6 GCGCGCCGCCATGCGAGCTGGAGAGCGGAGAACTCGCTCTTCTCGCACT 65
DB 1 GCGCGCCGCCATGCGAGCTGGAGAGCGGAGAACTCGCTCTTCTCGCACT 60
QY 66 CCGGAGCGAGGCTCCAGTCGGAGAGCGAAGATGTTCTCTCAAGATGGAACG 125
DB 61 CCGGAGCGAGGCTCCAGTCGGAGAGCGAAGATGTTCTCTCAAGATGGAACG 120
QY 126 CCGTAGCATGTGAGCTGGAGCTTGAGTCGATACCTGTGCACTTCGAGGTC 185
DB 121 CCGTAGCATGTGAGCTGGAGCTTGAGTCGATACCTGTGCACTTCGAGGTC 180
QY 186 TGATGATGCTGCTGCTGATGTCGAAGTGAACAAGCAGAGAGAGCTGTGGGCT 245
DB 181 TGATGATGCTGCTGCTGATGTCGAAGTGAACAAGCAGAGAGCTGTGGGCT 240
QY 246 GGGAGAGGTGTAACCATCTTCCACAACATGCTCATGTCCCTTGCGTGAACAGACA 305
DB 241 GGGAGAGGTGTAACCATCTTCCACAACATGCTCATGTCCCTTGCGTGAACAGACA 300

QY 306 ATGCGTCCCTCTGCTGGCAGAGACTGTGATGCCAAGAAATGGCAATGAGAGGTG 365
DB 301 ATGCGTCCCTCTGCTGGCAGAGACTGTGATGCCAAGAAATGAGAGGTG 360
QY 366 CCCAGGCGCTCCGT 425
DB 361 CCCAGGCGCTCCGT 420
QY 426 ATCCTTAGAGAGAGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 485
DB 421 ATCCTTAGAGAGAGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 480
QY 486 ATTTGCTGTATTTTGGGAATTCCTCAATTAAGTAATTTGTAATAATGGCCT 545
DB 481 ATTTGCTGTATTTTGGGAATTCCTCAATTAAGTAATTTGTAATAATGGCCT 540
QY 546 TTCCCTACCTGCTGT 605
DB 541 TTCCCTACCTGCTGT 600
QY 606 TGATCTTTGTTTATCTGTACCCAGCT--GGAACATGTGTTCACAGAAACATTTGT 662
DB 601 TGATCTTTGTTTATCTGTACCCAGCTGTGGAAACATTTGTTCACAGAAACATTTGT 660
QY 663 TTGTTTATGCTGTGAGGGTTAAAAAATAGATAAGCAATGTTACAGTAACAAATTAAT 722
DB 661 TTGTTTATGCTGTGAGGGTTAAAAAATAGATAAGCAATGTTACAGTAACAAATTAAT 720
QY 723 GCATTGAAAGCCGACCTCTCTTAATCTTTTGTGTGTGTGTGTGTGTGTGTGTGT 782
DB 721 GCATTGAAAGCCGACCTCTCTTAATCTTTTGTGTGTGTGTGTGTGTGTGTGTGT 780
QY 783 CCCGCTGTCTTCAATTTGTCTGTGAATGAGATTTTAACTGTCACTGAAAGCGGTA 842
DB 781 CCCGCTGTCTTCAATTTGTCTGTGAATGAGATTTTAACTGTCACTGAAAGCGGTA 840
QY 843 ACTGCGGTAACTGTAAATATGCGTAACTGTGCGGTAAAGCGCTTGTCTGTGACTT 902
DB 841 ACTGCGGTAACTGTAAATATGCGTAACTGTGCGGTAAAGCGCTTGTCTGTGACTT 900
QY 903 CTCATCTTTGACTGTGGCAGGAAGCTGATGTTTCAACCATTTAGTTCTTAAAGAACTG 962
DB 901 CTCATCTTTGACTGTGGCAGGAAGCTGATGTTTCAACCATTTAGTTCTTAAAGAACTG 960
QY 963 TTTTCTGT 1022
DB 961 TTTTCTGT 1018
QY 1023 GCGTACTAGTAACTGCAAGTTCATTAATGCAATGTTTAAATAATTTGCTTTG 1082
DB 1019 GCGTACTAGTAACTGCAAGTTCATTAATGCAATGTTTAAATAATTTGCTTTG 1077
QY 1083 AGTATTAAAGTTGATATATACCTTAAATCATTAACCTTA 1125
DB 1078 AGTATTAAAGTTGATATATACCTTAAATCATTAACCTTA 1120

RESULT 3
BC008627 816 bp mRNA linear PRI 12-JUL-2001
LOCUS BC008627
DEFINITION Homo sapiens, ring finger protein 7, clone MGC:17274 IMAGE:4177613,
mRNA, complete cds.
ACCESSION BC008627
VERSION BC008627.1 GI:14250388
KEYWORDS MGC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 816)
AUTHORS Strausberg, R.
TITLE Direct Submission

				Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov			
				Series: IRAL Plate: 21 Row: 0 Column: 13			
				This clone was selected for full length sequencing because it passed the following selection criteria: Genomescan gene prediction, Similarity but not identity to protein.			
FEATURES	location/Qualifiers						
source	1..822						
	/organism="Homo sapiens"						
	/db_xref="LocusID:9616"						
	/db_xref="taxon:9606"						
	/clone="MGC:14618 IMAGE:4069078"						
	/tissue.type="Kidney, hypernephroma"						
	/clone_lib="NH_MGC_58"						
	/lab_host="DH10B"						
	/note="Vector: pDNR-LIB"						
	27..368						
	/codon_start=1						
	/product="ring finger protein 7"						
	/protein_id="AAH05966.1"						
	/db_xref="GI:13543636"						
	/translation="MADVEDGEETFCALASHGSSGSKSGGDKMFSLKKNVAVAMSMDVECTCAICRYQVMDACLRCQAEKNROEDCVVWMECHNSFHNCCSLMVYKNNRCPICQODPVVVOIRIC"						
BASE COUNT	240 a	175 c	210 g	197 t			
ORIGIN							
Query Match	31.2%;	Score 355.4;	DB 9;	Length 822;			
Best Local Similarity	78.6%;	Pred. No. 1.9e-79;					
Matches	502;	Conservative	0;	Mismatches	101;	Indels	36;
					Gaps		5;
QY	1	GTTTCGCCGCCGCCGCATATGCGCGAGCGAGTGGAGAGCGCGCAGAGAACCTTCGCTCTTC	60				
Db	11	GCTCGCGCGCCGCCGCAATGCGCGAGCGTGAAGACGAGAAACCTCGCCCTGGCCTC	70				
QY	61	GCACTCGGGAGCGGAGGCTCCAAATGCGGGAGCGCAAAAGTCTCTCCCAAGAGT	120				
Db	71	TCACCTCCGGAGACTGAGCTCCAACTCCGGAGCGCACAAAGATGTTCTCCCTCAAGAA	130				
QY	121	GAACGCGGTAGCATATGTGGAGCTGGAGCGTTGAGTGCATATCTGTCATTCGAGGGT	180				
Db	131	GAACGCGGTCCGCAATGTGGAGCTGGAGCTGGAGTGGCGATACGTCCGCATTCGACGGT	190				
QY	181	CCAGGTATGATGCTCTCTCCCTTCATATGTCAAAGCTTAAACAAAGCAGAGAGCTGTGT	240				
Db	191	CCAGGTATGATGCTCTCTCTTAAATGTCAAGCTTAAACAAAGCAGAGAGCTGTGTGT	250				
QY	241	GGTCGGGGAGAGTGAACCAATCCCTTCACAAAGCTGCAGATGTCCTGGGGTAAACA	300				
Db	251	GGTCGGGGAGAAATGTATCAATCTCTTCACAAAGCTGCAGATGTCCTGGGGTAAACA	310				
QY	301	GAACATATGCTGCCCTCTGTGTCCAGCAGACTGGGTAGTCCAAAGATCGCAATGAGA	360				
Db	311	GAACATATGCTGCCCTCTCTGTGTCCAGCAGACTGGGTAGTCCAAAGATCGCAATGAGA	370				
QY	361	GGTGGCCAGCGCTCTGTGTGTGCTGTGAC-----	393				
Db	371	GTGTGAAGAGCTCTTATGCGCAGTGTTCAGAGCCCTGTGTGATCTTGTATTCAGTG	430				
QY	394	CCTGGCAAAAGACTTAAACATGACGGGATTCATCTTGAGAGAG-----AGAGATGTG	449				
Db	431	CCCTACAAAGGCTAGAACACTACAGGGATGATTAATCTTCAAAATAGAGACCGATGATCTG	490				
QY	450	TGCGGCTTTGAGACTCACCAAAAGCGCTTATTAATTTGTCTGTAGTTAGTTTGGAAA	509				
Db	491	TGTCCTTTGGAGACTCAACAAAGCCTTGTTAGCATTTTGCAGTTTA-TCTTCAGAAA	549				
QY	510	TTTCTCTACAAATT---AAATATATTTGTTAAAAATGGCCTTTCCTACCTCTGTGTGTG	566				
Db	550	TTTCTCTGGAATTAAAGATATTAATTAAGAGTGCTCTCTTCACTACCTCTGTGTGTGTG	609				

Db	610	TCGGCCAC-ACAGCTTAGAAGTGTCTATAAAAAAGGAAA	647
RESULT	6		
LOCUS	AC108040/c		
DEFINITION	AC108040	163521 bp	DNA
ACCESSION	AC108040		Library HTG 30-JAN-2002
VERSION	AC108040.2	GI:18425316	
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTLP.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
TITLE	1 (bases 1 to 163521)		
REFERENCE	Waterston, R.H.		
AUTHORS	The sequence of Homo sapiens clone		
TITLE	unpublished		
REFERENCE	2 (bases 1 to 163521)		
AUTHORS	Waterston, R.H.		
TITLE	Direct Submission		
JOURNAL	Submitted (24-JAN-2002) Genome Sequencing Center, Washington		
	University School of Medicine, 4444 Forest Park Parkway, St. Louis,		
	MO 63108, USA		
COMMENT	On Jan 30, 2002 this sequence version replaced g1:183008830.		

REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1 (bases 1 to 163757) Kaul,R.K., Olson,M.V., Raymond,C., Clendenning,J., Ivey,R.G. and Haugen,E.D.
TITLE	Direct Submission
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 163757)
AUTHORS	Kaul,R.K., Olson,M.V., Raymond,C., Clendenning,J., Ivey,R.G. and Haugen,E.D.
TITLE	Direct Submission
JOURNAL	Submitted (04-JUL-2001) Genome Center, University of Washington, Box 352145, Seattle, WA 98195, USA
COMMENT	On Jul 4, 2001 this sequence version replaced gi:8705848.

```

----- Genome Center
Center: University of Washington Genome Center
Center Code: UWGC
Web site: http://www.genome.washington.edu
Contact: uwgchgs@u.washington.edu
Drafting Center: BCM

----- Project Information
Center project name: chr-3
Center clone name: RP11-229A12 (bc0315)

----- Summary Statistics
Sequencing vector: plasmid; 50% of reads
Sequencing vector: unknown; 50% of reads
Chemistry: Dye-terminator Big Dye; 47% of reads
Chemistry: Dye-terminator EBV Dye; 52% of reads
Assembly program: Phrap; version 0.99019
Consensus quality: 162824 bases at least Q40
Consensus quality: 163281 bases at least Q30
Consensus quality: 163468 bases at least Q20
Insert size: 163557; sum-of-contigs
Quality coverage: 7.6x in Q20 bases; sum-of-contigs

-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 15071: contig of 15071 bp in length
* 15072 15171: gap of unknown length
* 15172 57886: contig of 42715 bp in length
* 57887 57986: gap of unknown length
* 57987 163757: contig of 105771 bp in length.
*
* Location/Qualifiers
* 1. 163757
* /organism="Homo sapiens"
* /db_xref="taxon:9606"
* /chromosome="3"
* /clone="RP11-229A12"
* /clone_lib="RRCI human BAC library 11"
* 1. 15071
* /note="assembly_name:Contig6"
* 15172. 57886
* /note="assembly_name:Contig7"
* 57987. 163757
* /note="assembly_name:Contig8"
* 69589. 69590
* /note="g11154843(gbj)J018291(TRN10)S13 excised from bc0315
* Contig8 ( 11602 . 12930 )
* TGAAGATCCCTCATTAATTTCCCAAGCGGTAACTGTGTGAATAAATTTGAGCTT
* ATGAGGTCCTCAGCCAGAGTAAGTCTCTTGTATTTGAGGACGAATGCCGCA
* AAATCTCATGCGCTAAGCAGATCTTGAGAGTACGTTCATTTGACATGCTTAC
* CTGGAATGCTGTGGCCCAACCTCTTTCTGACAGTGAAGCCGCCAAGCCAACATGTT
* AGTTGAACATCAGGCGCATTTAGACATGATATCAAAAGCCTCTGAGCTGCTGTT
* GCGTATGACCGTAGGCTTAATGCTGAGGACAGATTTTCAAGTCTGGAGAAGTTCTTC
* AATCTGATTCGCTGCAATAGATTTAAACAAGTTGTTGGCTGTGATTTACACA
* GGTAAATGACTGCTGCAACCATGCGCTGAGTCTGAGTATTTAGGTTAGTG
* ACGGTGATGACATGATGATCGGTGTGAGGCGCTCATATTTTGGCTTTGAGGAGACA

```

[illegible]


```

* 11894 12661: contig of 768 bp in length
* 12662 12761: gap of 100 bp
* 12762 13525: contig of 764 bp in length
* 13526 13625: gap of 100 bp
* 13626 14347: contig of 722 bp in length
* 14348 14447: gap of 100 bp
* 14448 15172: contig of 725 bp in length
* 15173 15272: gap of 100 bp
* 15273 16007: contig of 735 bp in length
* 16008 16107: gap of 100 bp
* 16108 16855: contig of 748 bp in length
* 16856 16955: gap of 100 bp
* 16956 17711: contig of 756 bp in length
* 17712 17811: gap of 100 bp
* 17812 18568: contig of 757 bp in length
* 18569 18668: gap of 100 bp
* 18669 19421: contig of 753 bp in length
* 19422 19521: gap of 100 bp
* 19522 20229: contig of 708 bp in length
* 20230 20329: gap of 100 bp
* 20330 21061: contig of 732 bp in length
* 21062 21161: gap of 100 bp
* 21162 21909: contig of 748 bp in length
* 21910 22009: gap of 100 bp
* 22010 22773: contig of 764 bp in length
* 22774 22873: gap of 100 bp
* 22874 23623: contig of 750 bp in length
* 23624 23723: gap of 100 bp
* 23724 24489: contig of 766 bp in length
* 24490 24589: gap of 100 bp
* 24590 25354: contig of 765 bp in length
* 25355 25454: gap of 100 bp
* 25455 26174: contig of 720 bp in length
* 26175 26274: gap of 100 bp
* 26275 27043: contig of 769 bp in length
* 27044 27143: gap of 100 bp
* 27144 27900: contig of 757 bp in length
* 27901 28000: gap of 100 bp
* 28001 28753: contig of 753 bp in length
* 28754 28853: gap of 100 bp
* 28854 29572: contig of 719 bp in length
* 29573 29672: gap of 100 bp
* 29673 30419: contig of 747 bp in length
* 30420 30519: gap of 100 bp
* 30520 31262: contig of 743 bp in length
* 31263 31362: gap of 100 bp
* 31363 32052: contig of 690 bp in length
* 32053 32152: gap of 100 bp
* 32153 32909: contig of 757 bp in length
* 32910 33009: gap of 100 bp
* 33010 33736: contig of 727 bp in length
* 33737 33836: gap of 100 bp
* 33837 34598: contig of 762 bp in length
* 34599 34698: gap of 100 bp
* 34699 35456: contig of 758 bp in length
* 35457 35556: gap of 100 bp
* 35557 36274: contig of 718 bp in length
* 36275 36374: gap of 100 bp
* 36375 37111: contig of 737 bp in length
* 37112 37211: gap of 100 bp
* 37212 37954: contig of 743 bp in length
* 37955 38054: gap of 100 bp
* 38055 38783: contig of 729 bp in length
* 38784 38883: gap of 100 bp
* 38884 39686: contig of 803 bp in length
* 39687 39786: gap of 100 bp
* 39787 40535: contig of 749 bp in length
* 40536 40635: gap of 100 bp
* 40636 41387: contig of 752 bp in length
* 41388 41487: gap of 100 bp
* 41488 42254: contig of 767 bp in length
* 42255 42354: gap of 100 bp
* 42355 43046: contig of 692 bp in length

```

```

* 43047 43146: gap of 100 bp
* 43147 43885: contig of 739 bp in length
* 43886 43985: gap of 100 bp
* 43986 44741: contig of 756 bp in length
* 44742 44841: gap of 100 bp
* 44842 45604: contig of 763 bp in length
* 45605 45704: gap of 100 bp
* 45705 46429: contig of 725 bp in length
* 46430 46529: gap of 100 bp
* 46530 47244: contig of 715 bp in length
* 47245 47344: gap of 100 bp
* 47345 48088: contig of 744 bp in length
* 48089 48188: gap of 100 bp
* 48189 48948: contig of 760 bp in length
* 48949 49048: gap of 100 bp
* 49049 49803: contig of 755 bp in length
* 49804 49903: gap of 100 bp
* 49904 50648: contig of 745 bp in length
* 50649 50748: gap of 100 bp
* 50749 51506: contig of 758 bp in length
* 51507 51606: gap of 100 bp
* 51607 52345: contig of 739 bp in length
* 52346 52445: gap of 100 bp
* 52446 53185: contig of 740 bp in length
* 53186 53285: gap of 100 bp
* 53286 54059: contig of 774 bp in length
* 54060 54159: gap of 100 bp
* 54160 54901: contig of 742 bp in length
* 54902 55001: gap of 100 bp
* 55002 55740: contig of 739 bp in length
* 55741 55840: gap of 100 bp
* 55841 56577: contig of 737 bp in length
* 56578 56677: gap of 100 bp
* 56678 57433: contig of 756 bp in length
* 57434 57533: gap of 100 bp
* 57534 58290: contig of 757 bp in length

```

Query Match 26.9% Score 306.6; DR 2; Length 59155;
 Best Local Similarity 77.5%; Pred. No. 1; le-66;
 Matches 424; Conservative 0; Mismatches 89; Indels 34; Gaps 3;

```

QY 62 CACGCCGGAGCGAGCGCTCCAGTCCAGTGGAGGCGACAAAGTGTCTCTCCAGAAAGTGG 121
DB 7636 CACTCCAGGAGCGAGCGCTCCAACTCCGGAGGCGACAAAGTGTCTCAGCAAGAGTGG 7695
QY 122 AACGGGTAGCCATGTGGAGCTGGAGCTGAGTGCATACCTCTGCGCATCTGAGGTC 181
DB 7696 AACGTGTGGCCATGTGGAGCTGGGCGTGGAGTGAATATGTGCGCATCTGCGAGTGC 7755
QY 182 CAGGTGATGATGCTGCTGCTCGTGCATGTCACAGTCAAGTCCCTGCTGCTGCTGCTG 241
DB 7756 CAGGTGATGATGCTGCTGCTCGTGCATGTCACAGTCAAGTCCCTGCTGCTGCTGCTG 7815
QY 242 GTCTGGGAGAGTGTAACTATTCCTTCACAACTGCTGCATGCTCCCTGCTGCTGCTGCT 301
DB 7816 GTCTGGGAGAGTGTAACTATTCCTTCGCAACTGCTGCATGCTCCCTGCTGCTGCTGCT 7875
QY 302 AACAAATCGCTGCCCTCTGTGCCAGAGAGCTGGGAGTGCAGAAATCGCAATGAGAG 361
DB 7876 AACAAATCGCTGCCCTCTGTGCCAGAGAGCTGGGAGTGCAGAAATCGCAATGAGAG 7935
QY 362 GTGGCCAGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 394
DB 7936 CGGTAGAGAGCTTCTAGTGTGAGTGTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7995
QY 395 CTGGCAAAAGCTAAACCTGAGGAGATTCCTTGAAGAGAGAGAGAGAGAGAGAGAGAG 450
DB 7996 CCTTCAAAAGCTAAGAACCTGAGGAGATTCCTTGAAGAGAGAGAGAGAGAGAGAGAG 8055
QY 451 GCGCCTTGAAGCTAACCAAGCTGCTTGAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 510
DB 8056 GGTCTTGGAGACATCAAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8115

```

Query Match	26.8%	Score 305.2	DB 9	Length 342
Best Local Similarity	93.3%	Pred. No. 8.9e-67		
Matches 319	Conservative 0	Mismatch 23	Indels 0	Gaps 0
QY 17	ATGCGCGCATGTGGAGGAGCGCGAGGAACCCATGCGCTTTCTTTCGCACATCCGGGAGCCGA	76		
Db 1	ATGGCGCGCATGTGAAGAGAGAGAAACATCGCGCCCTCTACACTCCGGGAGCTCA	60		
QY 77	GGCTCCAAAGTGGGGAGGGGACAAAGATGTTCTCTCAAGAAAGTGGACACCGGTACCATG	136		
Db 61	GGCTCAACGTGGGGAGGGGACAAAGTGTTCCTCCCAAGAAATGGAAACGGGTGCCATG	120		
QY 137	TGGAACCTGGGACGTGTGATGCGATACCTCTGTGCCATCTCAGAGGTCAGAGTGATGATGCC	196		
Db 121	TGGACCTGGGACGTGTGATGCGATACGTGCGCCATCTCGAGGCTCCAGGTATGATGATGCC	180		
QY 197	TGCCCTGCATGTCAAGCTGAAAACAAGCAAGAGGACTGTGTGTGTCTGGGGAAGATGT	256		
Db 181	TGCTTTAATGATCAAGCTGAAAACAAGCAAGAGGACTGTGTGTGTCTGGGGAAGAAATGT	240		

Query Match	26.6%	Score 303.6	DB 6	Length 342
Best Local Similarity	93.0%	Pred. No. 2.3e-66		
Matches 318: Conservative	0	Mismatches 24	Indels 0	Gaps 0
QY 17 ATGGCCGACGTGGAGGACGGAGGACCCCTGCTCTTCTCTTCGACATCCGCGAGGCGCA	76			
Db 1 ATGGCCGACGTGGAGGACGGAGGAGGAAACCTCGCCCTCGCCTCTCAGCTCCGGAGCTCA	60			
QY 77 GGCCTCAAGTCCGGGAGGCGACAGATGTTCTCTCAAGAACTGGAGACCGGCTAGCCATG	136			
Db 61 GGCCTCAAGCTCGGGAGGCGACAGATGTTCTCTCCCTCAAGAACTGGAGACCCGGTGGCCATG	120			
QY 137 TGGAGCTGGGACGTTGATGTCGATACCTCTGTGCTCATCTCGAGGCTCCAGTATGATGCC	196			
Db 121 TGGAGCTGGGACGTGGAGATGGATCGTACGTGGCCATCTCGAGGCTCCAGTATGATGCC	180			
QY 197 TGCCTTCGATGTCAGAGCTGAAACAAAGCAAGAGGACTGTTGTGTGTCGGGAGAGATGT	256			
Db 181 TGTCTTAAGATGTCAGAGCTGAAACAAACAAAGAGGACTGTGTGTGTCTGGGGAATAATGT	240			
QY 257 AACCATTTCCCTTCACAACTGCTGCATGTCCTGTGGGTGAAACAGAGCAATGCTGCCT	316			
Db 241 AATCATTTCCCTTCACAACTGCTGCATGTCCTGTGGGTGAAACAGAGCAATGCTGCCT	300			
QY 317 CTGTGCCAGCAGAGACTGGGTAGTCCAAAGAAATGGCAATGA	358			
Db 301 CTGTGCCAGCAGAGACTGGGTAGTCCAAAGAAATGGCAATGA	342			

RESULT	12
LOCUS	AC098496/c
DEFINITION	Rattus norvegicus clone CH230-23c21, *** SEQUENCING IN PROGRESS
ACCESSION	AC098496
VERSION	AC098496
KEYWORDS	HTG; HTGS_PHASE1.
SOURCE	Norway rat.
ORGANISM	Rattus norvegicus

REFERENCE
AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

1 (bases 1 to 170087)

Munzy,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amarantunga,H.C., Are,J.R., Banks,T., Barbata,J.,
Benton,J., Bimaga,K., Blankenburg,K., Bonin,D., Bouck,J.,
Bowie,S., Brieve,M., Brown,E., Brown,M., Bryant,N.P., Buhy,C.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Coyle,M.D., Dethorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Diaper,H.,
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
Foster,P., Prantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
Hollins,B., Homs,F., Howard,S., Huber,J., Huliy,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jollivet,S.,
Joudh,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Lounsged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,U., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M.,
Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwankwo,S.,
Ogoh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
Peety,J., Perez,L., Peters,L., Plickens,R., Plums,E., Pu,L.L.,
Quiles,S., Saverly,G., Scherer,S., Scott,G., Shen,H., Shostkari,N.,
Sisson,I., Sodergren,E., Sonalke,T., Sparks,A., Stanley,H.,
Stone,H., Sutton,A., Svatek,A., Tabor,P., Tameisa,A., Tameisa,K.,
Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,
Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R.,
Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
Washington,S., Williams,G., Williamson,A., Wleczko,R., Wooden,S.,
Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

Unpublished
2 (bases 1 to 170087)

Worley,K.C.

Direct Submission

Submitted (24-OCT-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Dec 20, 2001 this sequence version replaced gi:17062512.

Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

Project Information

Center project name: GHEX

Center clone name: CH230-23C21

Summary Statistics

Assembly program: Phrap; version 0.990329First call to
findPhrapList

Consensus quality: 130294 bases at least Q40
Consensus quality: 139538 bases at least Q30
Consensus quality: 146952 bases at least Q20

Estimated insert size: 136372; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-gel estimation
Quality coverage: 2.1x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/edenbank_draft_data.html).

NOTE: This is a 'working draft' sequence. It currently

* consists of 62 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1	6781:	contig of 6781 bp in length
*	6881:	gap of unknown length
*	11935:	contig of 5054 bp in length
*	11936:	gap of unknown length
*	12035:	gap of unknown length
*	12036:	contig of 4708 bp in length
*	16743:	gap of unknown length
*	16744:	contig of 5663 bp in length
*	22507:	gap of unknown length
*	22607:	contig of 7175 bp in length
*	29782:	gap of unknown length
*	29882:	contig of 5688 bp in length
*	35570:	gap of unknown length
*	35669:	gap of unknown length
*	38472:	contig of 2803 bp in length
*	38572:	gap of unknown length
*	44500:	contig of 5928 bp in length
*	44600:	gap of unknown length
*	49796:	contig of 5196 bp in length
*	49896:	gap of unknown length
*	51963:	contig of 2067 bp in length
*	52063:	gap of unknown length
*	54511:	contig of 2448 bp in length
*	54611:	gap of unknown length
*	58439:	contig of 3828 bp in length
*	58539:	gap of unknown length
*	62904:	contig of 4365 bp in length
*	63004:	gap of unknown length
*	63005:	contig of 2360 bp in length
*	63365:	gap of unknown length
*	65465:	contig of 3337 bp in length
*	68801:	gap of unknown length
*	68802:	contig of 2165 bp in length
*	71067:	gap of unknown length
*	71166:	contig of 3904 bp in length
*	75070:	gap of unknown length
*	75170:	contig of 3428 bp in length
*	75171:	gap of unknown length
*	78598:	gap of unknown length
*	82817:	contig of 4119 bp in length
*	82818:	gap of unknown length
*	82918:	contig of 3200 bp in length
*	86217:	gap of unknown length
*	86218:	contig of 2252 bp in length
*	88470:	gap of unknown length
*	88569:	gap of unknown length
*	90299:	contig of 1730 bp in length
*	90300:	gap of unknown length
*	90399:	gap of unknown length
*	92498:	contig of 2099 bp in length
*	92499:	gap of unknown length
*	92599:	contig of 3110 bp in length
*	95709:	gap of unknown length
*	95808:	contig of 2641 bp in length
*	98449:	gap of unknown length
*	98450:	contig of 2154 bp in length
*	98550:	gap of unknown length
*	100703:	contig of 2154 bp in length
*	100803:	gap of unknown length
*	100804:	contig of 2869 bp in length
*	103392:	gap of unknown length
*	103393:	contig of 3452 bp in length
*	103493:	gap of unknown length
*	106944:	contig of 2704 bp in length
*	107045:	gap of unknown length
*	109748:	contig of 2557 bp in length
*	109848:	gap of unknown length
*	112505:	contig of 3570 bp in length
*	112605:	gap of unknown length
*	116175:	contig of 1647 bp in length
*	116275:	gap of unknown length
*	117922:	contig of 1647 bp in length
*	117923:	gap of unknown length
*	118022:	contig of 2049 bp in length
*	118023:	gap of unknown length
*	120071:	contig of 2049 bp in length
*	120072:	gap of unknown length
*	120171:	gap of unknown length

```

* 120172 121899: contig of 1728 bp in length
* 121900 121999: gap of unknown length
* 122000 125200: contig of 3201 bp in length
* 125201 125300: gap of unknown length
* 125301 127224: contig of 1924 bp in length
* 127225 127324: gap of unknown length
* 127325 129155: contig of 1831 bp in length
* 129156 129255: gap of unknown length
* 129256 130948: contig of 1693 bp in length
* 130949 131048: gap of unknown length
* 131049 133234: contig of 2186 bp in length
* 133235 133334: gap of unknown length
* 133335 135257: contig of 1923 bp in length
* 135258 135357: gap of unknown length
* 135358 137309: contig of 1952 bp in length
* 137310 137409: gap of unknown length
* 137410 139006: contig of 1597 bp in length
* 139007 139106: gap of unknown length
* 139107 141230: contig of 2124 bp in length
* 141231 141330: gap of unknown length
* 141331 142977: contig of 1647 bp in length
* 142978 143077: gap of unknown length
* 143078 14363: contig of 1286 bp in length
* 144364 14463: gap of unknown length
* 14464 145811: contig of 1348 bp in length
* 145812 145911: gap of unknown length
* 145912 147757: contig of 1846 bp in length
* 147758 147857: gap of unknown length
* 147858 148900: contig of 1043 bp in length
* 148901 149000: gap of unknown length
* 151046 151046: contig of 2046 bp in length
* 151047 151146: gap of unknown length
* 151147 152783: contig of 1637 bp in length
* 152784 152883: gap of unknown length
* 152884 154473: contig of 1590 bp in length
* 154474 154573: gap of unknown length
* 154574 155649: contig of 1076 bp in length
* 155650 155749: gap of unknown length
* 155750 156940: contig of 1191 bp in length
* 156941 157040: gap of unknown length
* 157041 158097: contig of 1057 bp in length
* 158098 158197: gap of unknown length
* 158198 160281: contig of 1984 bp in length
* 160282 161529: gap of unknown length
* 161530 161629: contig of 1248 bp in length

```

Query Match 22.5% Score 256.4; DB 2; Length 170087;
 Best Local Similarity 64.8%; Pred. No. 7.3e-54;
 Matches 653; Conservative 0; Mismatches 242; Indels 112; Gaps 14;

```

OY 235 TGTGTGCTGTGGGAGAGTGAACCATTCCTTCACAACTGCTGCATGCTCCGTGGGT 294
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 48035 TCGAGTGGCTGTGGGAGAGTGAACCATTCCTTCACAACTGCTGCATGCTCCGTGGGT 47976

OY 295 GAAACAGAACATCGCTGCTGTGTCACAGAGACCTGGGTGTCACAAAGATGGGCA 354
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 47975 GAAACAGAACATCGCTGCTGTGTCACAGAGACCTGGGTGTCACAAAGATGGGCA 47916

OY 355 ATGAGAGTGGCCAGAGCGCTCTGTGTGTCAGACCTGGGACA----- 402
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 47915 GTGAGAGCGCGAGAGCGCTCTGTGTGTCAGACCTGGGACAACCTGTCACCAA 47856

OY 403 -----AGACTAAACATCGACAGGATTCCTTGAGAGAGAGAGATGCTGTGC 452
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 47855 GTGACTGAGACGACTTGACATCGACAGGATTCACCTTCAGATGAGAGATGCTGTGC 47795

OY 453 GCCTTGAGACTACCAAGAGCTTCTTATTAATTTGCTTTGTTAGTTTGGGAATTC 512
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 47796 GCCTTGAGAGCTCATCCAAAGCATGGTTA-TCATTTGCTCTTTTAAATTTTGGAAAAATG 47738

OY 513 TCTACAATTAAGATTAATTTGTTAAAAAGCCTTCTCACTCTGTGTGTGTGTGAT 572
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

DB 47737 TCTACAATTAAGATTAATTTAT-AAAATAGCCTTCTCACTCTGTGTGTGTGATAT 47679
OY 573 ACGAATGCATAGAGAGCGGAGAACACCAAAATGATCTTTGTTATCTGTACCA----- 628
DB 47678 GTACTGCTTTGAACACCGAATGCAGAGATCTTTGTGTATCTGTCACTGATATAGGG 47619
OY 629 CGACTGGAACATGTGTGTACAGAGACATTTGTTGTATTTGCTGTGAGGTTAAAA 688
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 47618 CTGGAGAGCATTTGCGTTTCACAGACATTTGGGT-TTATGCTTGGGGCTAAAAAA 47560
OY 689 ATAGATAAACGAATGTATACAGTAACAATAAATGATGAAAAAGCCACTCTCC- 744
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 47559 ATAAATAAAAAGATTTTACACTAAAAAATGATGAGAACCACTACCTACCT 47500
OY 745 -----TATCTTTTGTGGGAGAGGCAAGCAGCGCCACTGCTCTTCAT 797
DB 47499 TGTCTAATTAACCTTTGTGTGGAGTGAGCAACCAAGGTATCTTTTATATTA 47440
OY 798 TTG-----CTGTGAATGAGATTTTAACCTG 823
DB 47439 ATGAGGTTAATCTGTGACTGAGTAAGAGGTTAAATCCAGTGTATTAAGAAACAG 47380
OY 824 CACTCAGTGAAGAGCGCTAATCTGCGTAA-----CTG 858
DB 47379 AATCGGGCTGGGGATTTAGCTCAGTGTAGAGCGCTCTTGAAGAGCGCAAGCCCTG 47320
OY 859 TAATATGGCTAATCTCGGGTAAAGCGCTTGTCTCTGACTTCCATCTTTGACTTG 918
DB 47319 GGTTCGGTCCCGAGCTCCGAAAAAGACCAAAAAAAGAAAAAGAAACCGAAG 47260
OY 919 GCCAGAGAC-----CTGATTTGTCACCACTAATGCTTAAAGACTGTTTCTGTTT 973
DB 47259 TCAAGGAACACTGTTTCACTGTTCACCGCGCTTAAAGGTTAAAGATTTTCTGTTTC 47200
OY 974 TCGCGAAGTGTATTGTATGTTTAAAGAAAAATATAGTGAAGAAATGCGTTACTAGT 1033
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 47199 TG-GAAGGTGATTTGATTTAGTTC-AAAATATGATGGGAAAT-GCTTACAGT 47143
OY 1034 ATACACTGAAGTCAATATGCAATGTTTAA-----AAATTTGCTTTAGTATTT 1089
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 47142 ATGGCACCGAGTCAATTAAGCAATGCTTAAATGAATATTTGCTTGAATTAAT 47083
OY 1090 AAGTTGATATATCTTAAATCATTAATCAATTAATCAATTA 1136
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 47082 AAGTTGATCTCTACTCTTAATTAATTAATTAATCAATTAATCAATTAAT 47036

```

RESULT 13
 AC060766/c 169675 bp DNA linear PRI 01-DEC-2001
 LOCUS AC060766 Homo sapiens chromosome 17, clone RP11-686D22, complete sequence.
 DEFINITION AC060766 GI:16931016
 ACCESSION AC060766
 VERSION AC060766.8
 KEYWORDS HTG.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 169675)
 Birren,B., Linton,L., Nussbaum,C., Lander,E., Abraham,H., Allen,N.,
 Anderson,S., Baldwin,J., Barne,N., Bastien,V., Bede,F.,
 Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G.,
 Campopiano,A., Castle,A., Choepell,Y., Colangelo,M., Collins,S.,
 Collamore,A., Cooke,P., Dearellano,K., Delaney,K., Diaz,J.S.,
 Dodge,S., Domino,M., Doyle,M., Ferrelita,P., Fitzhugh,W., Gage,D.,
 Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
 Grand-Pierre,N., Grant,G., Hagos,B., Heatford,A., Horton,L.,
 Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
 Klein,J., Lacroque,K., Lamazares,R., Landers,T., Lehoczy,J.,


```

repeat__region    /rpt_family="Alusq"
                    /note="less than 30 qual SNGI region"
repeat__region    /rpt_family="LI"
                    complement(24452,.24451)
repeat__region    /rpt_family="LIMEC"
                    complement(25527,.25159)
repeat__region    /rpt_family="LIMEC"
                    complement(25527,.2528)
repeat__region    /rpt_family="HERVJ4"
                    25739,.25762
repeat__region    /rpt_family="AT_rich"
                    27209,.27454
repeat__region    /rpt_family="HALIB"
                    complement(27356,.27359)
unsure            /note="less than 30 qual SNGI region"
unsure            complement(27485,.27493)
unsure            /note="less than 30 qual SNGI region"
                    complement(28016,.28052)
repeat__region    /rpt_family="MER74B"
                    complement(28110,.28337)
repeat__region    /rpt_family="HERVL74"
                    complement(28533,.28596)
repeat__region    /rpt_family="HERVL74"
                    complement(28598,.28676)
repeat__region    /rpt_family="AUSJO"
                    complement(28877,.29267)
repeat__region    /rpt_family="HERVL74"
                    complement(29268,.29553)
repeat__region    /rpt_family="Alusq"
                    complement(29554,.29781)

```

Query Match	20.88;	Score 236.6;	DB 9;	Length 169675;
Best Local Similarity	68.58;	Pred. No. 7.8e-49;		
Matches 416;	Conservative 0;	Mismatches 154;	Indels 37;	Gaps 5

QY	1	GTCGCGCCCGCCGCAATGGCGAGCTGTGAGGAGCGGAGGAAACCTGGTCTTCTTC	60
Db	93487	GCTGTGCACAGCTGCCAATGCTAACGTGTGAGATGCGACAGAGGCCCTTGCTTCTC	934282
QY	61	GCACTCCGGGAGCGCAGGCTCCAACTCGGAGGCGA--CAAGATGTCTCTCAAGAA	118
Db	93427	TCACCTGTGGAGACACAGGCTCCAAATTGGAAGGGCGACAAAAAATGTCTCTCTCCAGGG	93368
QY	119	TGGAACGGGATAGGCATGTGAGAGCTGGGAGACGTGAGCTCGAATCCTGTGCCATCTGCAG	178
Db	93367	TGAACACAGTGGGATGTAGAGCTGCGCACACAGGAGTCCCATCAACATGCCATCTGCAG	93308
QY	179	GTCGAGGTGATGATGATCGCTGCTTGTGATGTCAAGTCAAGTGAACACAGCAGAGGAGCTGT	238
Db	93307	GTTGAGGTGATGATGATGCTTTCTTGTGGGTGTCAGCTGAAAA---CAAGAGAACTGTGCT	93255
QY	239	GTGCTCTGGGAGAGATGTAACCATTCCTTCCACAACCTGCTGATGCTCCCTGTGGTGAAA	298
Db	93251	GTCATCTGAGAGAGATGTAATCACCTCTTCCACAACCTGTGATGCTCCTGGGTACAA	93192
QY	299	CAGAACATCGCTGCGCTGCTGTGCCAGGAGAGCTGGTATGTCACAAAGATCGGCAATGA	358
Db	93191	CAGAGCATCTACCTGACTCTCTGCGCAGCAGACTGGATGGTCTTAAGATTAAGAGACGGA	93132
QY	359	GAGGTGGCCAGGCGCTCTGGTGTGGTCT-----G	391
Db	93131	GAGTAGCGAAGGCTTTCTTAGTGACCTTGGCCCAACCTTGGTGGATCTGTATCAAG	93072
QY	392	ACCCGCGACAAAGATTAACACTGAGGAGGATTCATCTGTGAGA--GAGAGATGCTGT	450
Db	93071	TGTCCACAAAGAGCTGGAACATCCAGGGGATTAATTTCTCAATAGGACACCAATGAATC	93012
QY	451	GCGCTTTGAGACTCACCAAAGGCTGTGTTATTTATTTATTTGCTTTATGTTTGGGAAT	510
Db	93011	TGTCCTTTGGAACATCACAAGCGAATTTTCAATTTTCATCAGTTTATTTTCAGGAAT	92955
QY	511	TCTCTCAATT---AAGATATTTTGTAAATAATGGCCTTTCTTACTCTGTGTGTGTGT	567

Db 92951 TCTCTACCAATTTAAGAAGATTAATTCTATAAAACAGCTTTTCTACTCTGAGGTGTGTGT 92892

QY 568 GTGATAC 574

Db 92891 CACACAC 92885

RESULT	14
AC106176/c	
LOCUS	160012 bp DNA linear HTG-12-JAN-2002
DEFINITION	Rattus norvegicus clone C8230-20P16, *** SEQUENCING IN PROGRESS
ACCESSION	AC106176 ***, 77 unordered pieces.
VERSION	AC106176
KEYWORDS	AC106176.1 GI:18138697
SOURCE	HTG; HTGS_PHASE1.
REMARKS	Norway rat. ..

REFERENCE AUTHORS

REFERENCE
AUTHORS

1 (bases 1 to 160012)

Muny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C., Alsbrooks, S.L., Amarantunge, H.C., Atre, J.R., Banks, T., Barbata, J., Benton, J., Bismare, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowie, S., Brieve, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Butch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Caron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhury, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Day, Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, G., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hoques, M., Holloway, C., Hollins, B., Homs, F., Howard, S., Huber, J., Hult, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolyet, S., Joudan, S., Karlsson, E., Kelly, S., Khan, U., King, L., Koryak, J., Kovar, C., Kralovic, J., Kuresh, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Licharge, O., Lieu, C., Liu, J., Liu, W., Louisgled, H., Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mel, G., Metzner, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J.K., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nguyen, E., Nockenkov, S., Ogih, M., Okunoye, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojchokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shoochari, N., Sisson, I., Sodergren, E., Sonake, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Tellro, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, Y., Villalón, D., Vinson, R., Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Wellington, S., Williams, G., Williamson, J., Wleczyk, R., Woden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstein, G., and Gibbs, R.

Direct Submission

Unpublished

2 (bases 1 to 160012)

Worley, K.C.

Direct Submission

Submitted (12-JAN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

Genome Center
Center: Baylor College of Medicine
Web site: <http://www.hgsc.bcm.tmc.edu/>

COMMENT

Db 23953 GCATTCTGGGAGCCAGAGCTCCAGATGCGGAGGGGACACAGATGTTCTCCCTCAGAGAGTG 23894
QY 121 GAACGGCGGTAGCCATGTGGAGCTGGGACGTTGAGTCGATACCTGTGCACTGTCAGAGGT 180
Db 23893 GAACGGCGGTAGCCATGTGGAGCTGGGACGTTGAGTCGATACCTGTGCACTGTCAGAGGT 23834
QY 181 CCAGGTGATGGAT 193
Db 23833 CCAGGTGATGGAT 23821

RESULT 15
HSA323208/c 439 bp DNA linear PRI 01-OCT-2001
LOCUS HSA323208 Homo sapiens genomic sequence surrounding NotI site, clone
DEFINITION NBI-073R.
ACCESSION AJ323208
VERSION AJ323208.1 GI:15867587
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 439)
Kutsenko, A.S., Gizatullin, R.Z., Al-Amin, A.N., Wang, F.,
Podowski, R.M., Matushkin, Y.G., Kvasha, S.M., Gyanchandani, A.,
Muravenko, O.V., Protodopov, A.I., Kashuba, V.I., Kiselev, L.D.,
Wasserman, W., Mahlestedt, C. and Zabarovsky, E.R.
Analysis of NotI flanking sequences: a new tool for gene discovery
and verification of the human genome
Unpublished
JOURNAL 2 (bases 1 to 439)
REFERENCE Zabarovsky, E.R.
AUTHORS Direct Submission
TITLE Submitted (16-MAY-2001) Microbiology and Tumorbiology Centre,
JOURNAL Karolinska Institute, Theorells vag, 3, Box 280, Stockholm 171 77,
Sweden

FEATURES
source location/Qualifiers
1..439
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NBI-073R"

BASE COUNT 63 a 161 c 146 g 68 t 1 others
ORIGIN

Query Match 14.4%; Score 164.2; DB 9; Length 439;
Best local Similarity 90.7%; Pred. No. 5.8e-31;
Matches 175; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 GTTCTGCGCGCGCCGACATGCGCGACGCTGGAGAGCGGAGGACCCCTGCTTCTTC 60
Db 383 GCTCGCGCGCGCGCCGACATGCGCGACGCTGGAGAGGAGGAAACCTGCGCCCTGCGCTC 324
QY 61 GCACTCCGGAGCGCAGAGCTCCAAGTGGGAGGCGGACACAAGATGTTCTCTCTCAAGAGTG 120
Db 323 TCACTCCGGAGCGCAGAGCTCCAAGTGGGAGGCGGACACAAGATGTTCTCTCTCAAGAGTG 264
QY 121 GAACGGCGGTAGCCATGTGGAGCTGGGACGTTGAGTCGATACCTGTGCACTGTCAGAGGT 180
Db 263 GAACGGCGGTAGCCATGTGGAGCTGGGACGTTGAGTCGATACCTGTGCACTGTCAGAGGT 204
QY 181 CCAGGTGATGGAT 193
Db 203 CCAGGTGATGGAT 191

Search completed: July 25, 2002, 19:10:29
Job time: 18108 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 25, 2002, 19:07:32 ; Search time 102.69 Seconds
(without alignments)
1803.561 Million cell updates/sec

Title: US-09-509-779-3

Perfect score: 754
Sequence: 1 ATGGCCGACGCGAAGACG.....CAATTAAGTCAGTTTAA 754

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 38353 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents_NA: *
1: /cgn2_6/ptodata/2/lna/5A.COMB.seq: *
2: /cgn2_6/ptodata/2/lna/5B.COMB.seq: *
3: /cgn2_6/ptodata/2/lna/6A.COMB.seq: *
4: /cgn2_6/ptodata/2/lna/6B.COMB.seq: *
5: /cgn2_6/ptodata/2/lna/PCTUS.COMB.seq: *
6: /cgn2_6/ptodata/2/lna/Backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	36	4.8	1342	3	US-08-832-399-1
C 2	36	4.8	1342	4	US-09-372-498-1
C 3	36	4.8	1575	3	US-08-858-876A-1
C 4	36	4.8	1575	4	US-09-472-880-1
5	35.6	4.7	176373	3	US-09-128-155-17
6	35.2	4.7	716	4	US-08-998-416-746
7	34.2	4.5	433	4	US-09-040-984-58
8	34.2	4.5	433	4	US-09-123-912-58
C 9	33.2	4.4	302	4	US-09-656-286-10
C 10	33.2	4.4	302	4	US-09-656-286-12
C 11	33.2	4.4	423	1	US-08-470-179-55
C 12	32.2	4.3	1779	4	US-09-371-696-1
C 13	31.6	4.2	302	4	US-09-656-286-11
C 14	31.6	4.2	302	4	US-09-656-286-13
C 15	31.4	4.2	423	1	US-08-470-179-142
C 16	31.4	4.2	3404	1	US-08-265-429A-1
17	31.4	4.2	3404	5	PCT-US95-09069-1
C 18	31.4	4.2	7218	1	US-08-232-463-14
C 19	31.2	4.1	1631	3	US-09-118-319-1
20	31	4.1	4195	3	US-08-340-011-1
21	31	4.1	4195	3	US-08-901-710-1
22	31	4.1	4416	3	US-08-795-430-1
23	31	4.1	4425	1	US-08-222-616-31
24	31	4.1	4425	5	PCT-US95-04228-31
25	31	4.1	4795	1	US-08-340-011-3
26	31	4.1	4795	3	US-08-901-710-3
27	31	4.1	9108	5	PCT-US95-04228-45

C 28	31	4.1	35081	2	US-08-752-760A-1	Sequence 1, Appl
C 29	30.8	4.1	1903	3	US-08-581-148C-15	Sequence 15, Appl
C 30	30.8	4.1	6924	1	US-08-015-973-2	Sequence 2, Appl
31	30.8	4.1	6924	2	US-08-448-164-2	Sequence 20, Appl
32	30.6	4.1	246240	2	US-08-724-394A-20	Sequence 21, Appl
33	30.6	4.1	246240	2	US-08-724-394A-21	Sequence 22, Appl
34	30.6	4.1	246240	2	US-08-724-394A-22	Sequence 3, Appl
C 35	30.4	4.0	1185	3	US-09-023-339-3	Sequence 5, Appl
C 36	30.4	4.0	1260	3	US-09-023-173-5	Sequence 10, Appl
C 37	30.4	4.0	1308	3	US-09-023-339-6	Sequence 6, Appl
C 38	30.2	4.0	791	4	US-08-858-207A-219	Sequence 19, Appl
39	30.2	4.0	4451	3	US-08-717-294-42	Sequence 42, Appl
40	30.2	4.0	4451	3	US-09-154-083-5	Sequence 5, Appl
C 41	30	4.0	302	4	US-09-656-286-14	Sequence 14, Appl
C 42	30	4.0	302	4	US-09-656-286-15	Sequence 15, Appl
C 43	30	4.0	302	4	US-08-644-326-1	Sequence 1, Appl
C 44	30	4.0	400	1	US-09-022-238-1	Sequence 1, Appl
C 45	30	4.0	400	4	US-09-022-238-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1

US-08-832-399-1/c
; Sequence 1, Application US/08832399

; Patent No. 6008050

; GENERAL INFORMATION:

; APPLICANT: Bergsma, Dirk

; APPLICANT: Shabon, Usman

; TITLE OF INVENTION: NOVEL HUMAN NEURENIN RECEPTOR TYPE 2

; NUMBER OF SEQUENCES: 12

; CORRESPONDENCE ADDRESS:

ADDRESSEE: SmithKline Beecham Corporation

STREET: 709 Swedeland Road

CITY: King of Prussia

STATE: PA

COUNTRY:

ZIP: 19406

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/832,399

FILING DATE: 02-Apr-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: King, William T

REGISTRATION NUMBER: 30,954

REFERENCE/DOCKET NUMBER: GH50020

TELECOMMUNICATION INFORMATION:

TELEPHONE: 610-270-5515

TELEFAX: 610-270-5090

TELEX:

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1342 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

US-08-832-399-1

Query Match 4.8%; Score 36; DB 3; Length 1342;
Best Local Similarity 54.5%; Pred. No. 0.11;
Matches 72; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY	28	ACCTGCGGCCGCTCTCATCTCGGGAGCTCAGGCTCCAAAGTCGGGAGGCGACAAAGATG	87
Db	645	ACCAAGACCGGTGCACACTCGGAGGGGGGCTCGGGCTCCCGCTTCGCGCTTCAGATTGC	586
QY	88	TTTCCTCCCTCAAGAGTAGTGAAGCGCGGTGGCCATGTGTGAGCTGTGGACGTGGATCGGATTACG	147
Db	585	TGCTTCTTCGCCCATGATGACGCGCATGTGGGCGGGCGAGGCCGACGAGCGGCCCCACGAG	526
QY	148	TGCGCCATCTGC	159
Db	525	AGCGCCACGACG	514

```

RESULT 2
US-09-372-498-1/c
: Sequence 1, Application US/09372498
: Patent No. 6166182
: GENERAL INFORMATION:
: APPLICANT: Derk J. Bergsma
: APPLICANT: Usman Shabon
: TITLE OF INVENTION: NOVEL HUMAN NEURENENSIN RECEPTOR TYPE 2
: TITLE OF INVENTION: AND SPICE VARIANTS THEREOF
: FILE REFERENCE: GH-50020-1
: CURRENT APPLICATION NUMBER: US/09/372,498
: CURRENT FILING DATE: 1999-08-11
: PRIOR APPLICATION NUMBER: 08/832,399
: PRIOR FILING DATE: 1997-04-02
: NUMBER OF SEQ ID NOS: 12
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 1
: LENGTH: 1342
: TYPE: DNA
: ORGANISM: HOMO SAPIENS
: US-09-372-498-1

```

Query Match	4.8%	Score 36	DB 4	Length 1342
Best Local Similarity	54.5%	Pred. No. 0.11		
Matches 72	Conservative	0	Mismatches 60	Indels 0
			Gaps	
QY	28	ACCTGCGCCCTTGAGCTTCACATCCGGAGGCTCCAAAGTCGGGAGGCGCAAGATG	87	
Db	645	ACGAGCAACGGTGACATCCGGAGGCGGGGCTCGGGCTCCCGCTCTCGAATTGC	586	
QY	88	TTTCCTCCCTCAGAGAGTGGAAACGGGTGGCCATGTGGAGCTGGGACCTGGAGTGCATACG	147	
Db	585	TGCTTTCGCCCCATGAGAGCGGCATGGGCGAGGGCGCGGACGAGCGGCGCCACGAG	526	
QY	148	TGGGCAATCTGC	159	
Db	525	AGCGCCACCGAC	514	

RESULT 3
 US-08-858-876A-1/c
 Sequence 1, Application US/08858876A
 Patent No. 6022856
 GENERAL INFORMATION:
 APPLICANT: Daniel CAPUT
 APPLICANT: Pascale CHALON
 APPLICANT: Pascual FERRARA
 APPLICANT: VITA NAVARRO
 TITLE OF INVENTION: TYPE 2 Neurotensin Receptor
 TITLE OF INVENTION: (INT-R2)
 NUMBER OF SEQUENCES: 12
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Jacobson, Price, Holman & Stern, PLLC
 STREET: 400 Seventh Street
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20004
 COMPUTER READABLE FORM.

```

1  MEDIUM TYPE: Floppy disk
2  COMPUTER: IBM PC compatible
3  OPERATING SYSTEM: PC-DOS/MS-DOS
4  SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
5
6  CURRENT APPLICATION DATA:
7  APPLICATION NUMBER: US/08/858,876A
8  FILING DATE: 19-SEP-1997
9  CLASSIFICATION: 536
10
11 PRIOR APPLICATION DATA:
12 APPLICATION NUMBER: PCT/FR 9723204
13 FILING DATE: 17-MAR-1997
14 ATTORNEY/AGENT INFORMATION:
15 NAME: Player, William E.
16 REGISTRATION NUMBER: 31,049
17
18 INFORMATION FOR SEQ ID NO: 1:
19
20 SEQUENCE CHARACTERISTICS:
21 LENGTH: 1575 base pairs
22 TYPE: nucleic acid
23 STRANDEDNESS: single
24 TOPOLOGY: linear
25
26 MOLECULE TYPE: cDNA
27 ORIGINAL SOURCE:
28 ORGANISM: Homo sapiens
29
30 FEATURE:
31 NAME/KEY: CDS
32 LOCATION: 37..1266
33
34 US-08-858-876A-1

```

Query Match	4.8%	Score 36;	DB 3;	Length 1575;
Best Local Similarity	54.5%	Pred. No. 0.12;		
Matches 72;	Conservative 0;	Mismatches 60;	Indels 0;	Gaps 0;

[illegible]

RESULT 4
US-09-472-880-1/c
Sequence 1, Application US/09472880
Patent No. 6274333
GENERAL INFORMATION:
APPLICANT: Daniel CAPUT
Pascale CHALON
Pascale FERRARA
Vita NATALIO
TITLE OF INVENTION: Type 2 Neurotensin Receptor
(hMT-R2)
NUMBER OF SEQUENCES: 12
CORRESPONDING ADDRESS:
ADDRESSEE: Jacobson, Price, Holman & Stern, PLLC
STREET: 400 Seventh Street
CITY: Washington
STATE: D. C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION NUMBER: US/09/472,880
FILING DATE: 28-Dec-1999

```
CLASSIFICATION: <Unknown>
PRIORITY: 1
APPLICATION NUMBER: PCT/FR 9723204
FILING DATE: 17-MAR-1997
ATTORNEY/AGENT INFORMATION:
  NAME: Player, William E.
  REGISTRATION NUMBER: 31,049
INFORMATION FOR SEQ ID NO: 1
SEQUENCE CHARACTERISTICS:
  LENGTH: 1575 base pairs
  TYPE: nucleic acid
  STRANDEDNESS: single
  TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
  ORGANISM: Homo sapiens
FEATURE:
  NAME/KEY: CDS
  LOCATION: 37..1266
  SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-472-880-1

Query Match
Best Local Similarity 4.8%; Score 36; DB 4; Length 1575;
Matches 72; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 28 ACCTGGCCCTGCTCTACCTCCAGCTCCAGCGGAGGCGCAGATG 87
DB 629 ACCAGACCGTCGACACGCGGCGGCTCCCGCTCCCGCTCGAGTTGC 570
QY 88 TTCTCCCTCAAGAGTGAACGCGGTGGCCATGTGAGCTGGAGCTGAGTGCATAG 147
DB 569 TGCTTCTGCGCCCATGATGACGGCCATGGGAGGCGAGGCGGCCACGAG 510
QY 148 TGGCCCATCTGC 159
DB 509 AGCGCCACGACG 498

RESULT 5
US-09-128-155-17
Sequence 17, Application US/09128155
Patent No. 6117654
GENERAL INFORMATION:
  APPLICANT: Pan, Yang
  TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY
  TITLE OF INVENTION: AND USES THEREOF
  FILE REFERENCE: 09404/052001
  CURRENT APPLICATION NUMBER: US/09/128,155
  CURRENT FILING DATE: 1998-08-03
  EARLIER APPLICATION NUMBER: US 60/091,650
  EARLIER FILING DATE: 1998-07-02
  EARLIER APPLICATION NUMBER: US 60/054,646
  EARLIER FILING DATE: 1997-08-04
  NUMBER OF SEQ ID NOS: 18
  SOFTWARE: FastSeq for Windows Version 3.0
  SEQ ID NO: 17
  LENGTH: 176373
  TYPE: DNA
  ORGANISM: Homo sapiens
  FEATURE:
    NAME/KEY: misc_feature
    LOCATION: (1)..(176373)
    OTHER INFORMATION: n = A,T,C or G
US-09-128-155-17

Query Match
Best Local Similarity 4.7%; Score 35.6; DB 3; Length 176373;
Matches 89; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 367 CAGTTGTCAGAGCCCTGGTGCATCTTGAATCCAGTCCCTCAAGGCTAGACACATCA 426
```

```
DB 164426 caggtgtgtgactcctagtcacagagttcttcttgcactgcgcccgaggtgtctaactcta 164485
QY 427 CAGGGATGATATTTCTTCATATAGACCCGATGATCTGTGGTCTTGCATCTCAAGC 486
DB 164486 ctgtgtcttaagactactcttggaggtcttcttaaaaaagagatctcacacactgagatt 164545
QY 487 CTTGTTAGCATTTGTCAGTTTATCTTCAGAAATTCCTGTCGATTAAGAGATAATT 544
DB 164546 ctgtgttaactcttlttggatgtagctcagggatctagcttgcttcttaaaaaaaaaact 164603

RESULT 6
US-08-998-416-746
Sequence 746, Application US/08998416
Patent No. 6239264
GENERAL INFORMATION:
  APPLICANT: Philippsen, Peter
  APPLICANT: Pohlmann, Rainer
  APPLICANT: Steiner, Sabine
  APPLICANT: Mohr, Christine
  APPLICANT: Wendland, Jurgen
  APPLICANT: Knechtle, Philipp
  TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSYPPII
  TITLE OF INVENTION: AND USES THEREOF
  NUMBER OF SEQUENCES: 1152
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: No. 6239264artis Corporation
    STREET: 3054 Cornwallis Road
    CITY: Research Triangle Park
    STATE: No. 6239264th Carolina
    COUNTRY: USA
    ZIP: 27709
COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: Patentin Release #1.0, Version #1.30
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/998,416
    FILING DATE: 24-DEC-1997
    CLASSIFICATION: 435
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER: CH 0016/97
      FILING DATE: 31-DEC-1996
      ATTORNEY/AGENT INFORMATION:
        NAME: Meigs, J. Timothy
        REGISTRATION NUMBER: 38,241
        REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
        TELECOMMUNICATION INFORMATION:
          TELEPHONE: 919-541-8587
          TELEFAX: 919-541-8689
        INFORMATION FOR SEQ ID NO: 746:
          SEQUENCE CHARACTERISTICS:
            LENGTH: 716 base pairs
            TYPE: nucleic acid
            STRANDEDNESS: single
            TOPOLOGY: linear
            MOLECULE TYPE: DNA (genomic)
            ORIGINAL SOURCE:
              ORGANISM: PAG1484UP
US-08-998-416-746

Query Match
Best Local Similarity 4.7%; Score 35.2; DB 4; Length 716;
Matches 79; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 56 GCTCAGGCTTCAAGTGGAGGCGACATGTTCTCCCTCAAGAGTGAACGCCGTGG 115
DB 375 GCGGGGGGTTCGAACGTGTTCAGAGTGTTCACGACATGTGAGAGCGCCACATCTCGT 434
```


	Query Match	4.28;	Score 31.6;	DB 4;	Length 302;
	Best Local Similarity	62.88;	Pred. No. 1.2;		
	Matches 49;	Conservative	0;	Mismatches 29;	Indels 0;
0y	227	TCGCGGAGCATGTAATCATCTTCCTCCACCAACGCTGCATGTCCTGGGGAAGACAGA	286		

Search completed: July 25, 2002, 19:08:16
Job time: 14187 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 25, 2002, 19:10:29 ; Search time 3583.2 Seconds
(without alignments)
4403.496 Million cell updates/sec

Title: US-09-509-779-3

Perfect score: 754
Sequence: 1 ATGGCCGACGTGAGACG.....CAATTAAGTGCAGTTTAA 754

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

```

GenEmbl:
1: gb.ba:*
2: gb.htg:*
3: gb.in:*
4: gb.om:*
5: gb.ov:*
6: gb.pat:*
7: gb.ph:*
8: gb.pl:*
9: gb.pr:*
10: gb.ro:*
11: gb.sts:*
12: gb.sy:*
13: gb.un:*
14: gb.vl:*
15: em.ba:*
16: em.fun:*
17: em.hum:*
18: em.in:*
19: em.mu:*
20: em.om:*
21: em.or:*
22: em.ov:*
23: em.pat:*
24: em.ph:*
25: em.pl:*
26: em.ro:*
27: em.sts:*
28: em.un:*
29: em.vl:*
30: em.htg.hum:*
31: em.htg.inv:*
32: em.htg.other:*
33: em.htg.inv:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	754	100.0	754	9	AF092878
2	691	91.6	836	9	AF164679
3	689.4	91.4	816	9	BC008627
4	663.6	88.0	822	9	BC005966
5	607.2	80.5	163521	2	AC108040
6	607.2	80.5	163757	2	AC092418
7	553.8	73.4	59155	2	AC104983
8	467.4	62.0	125041	2	AC068693
9	455.4	60.4	179878	2	AC092978
10	455.4	60.4	191221	2	AC067952
11	410.8	54.5	169675	9	AC060766
12	338.8	44.9	342	9	AF142060
13	337.2	44.7	342	6	AX282592
14	334	44.3	1140	10	AF092877
15	333.7	44.2	1125	10	BC011127
16	315.8	41.9	389	6	AX192664
17	314.4	41.7	388	6	AX192658
18	235	31.2	674	9	AF312226
19	175.4	23.3	439	9	HSA323208
20	175.4	23.3	191221	2	AC067952
21	154.6	20.5	160012	2	AC106176
22	154.6	20.5	170087	2	AC098496
23	99.2	13.2	125041	2	AC068693
24	84.6	11.2	166545	2	AC097192
25	83	11.0	170869	2	AC020266
26	83	11.0	170869	2	AC011696
27	83	11.0	171831	3	AC007473
28	83	11.0	278196	3	AE003825
29	80.4	10.7	219912	2	AC020855
30	77.8	10.3	5347	6	AX281690
31	76.8	10.2	554	9	BC017370
32	76	10.1	327	9	AF142059
33	76	10.1	497	9	HUMY060A05
34	76	10.1	508	9	AF140598
35	76	10.1	544	9	BC001466
36	75.2	10.0	504	10	AF140599
37	72.6	9.6	3484	10	HSRSTR
38	72.4	9.6	553	3	AY070810
39	72.4	9.6	750	3	AF218290
40	72.4	9.6	3185	3	AY061302
41	72.4	9.6	95294	2	AC019742
42	72.4	9.6	167962	3	AC010010
43	72.4	9.6	310120	3	AE003468
44	70.6	9.4	3208	6	AX212267
45	68.8	9.1	10899	3	CELRI0A10

ALIGNMENTS

RESULT 1
AF092878 754 bp mRNA linear PRI 24-JUL-2001
LOCUS Homo sapiens zinc RING finger protein SAG mRNA, complete cds.
DEFINITION AF092878
ACCESSION AF092878.1 GI:4588033
VERSION
KEYWORDS
SOURCE
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 754)
AUTHORS Swarcop,M., Bian,J., Aviram,M., Duan,H., Bisgaler,C.L., Loo,J.A.
and Sun,Y.
TITLE Expression, purification, and biochemical characterization of SAG,
a RING finger redox-sensitive protein
JOURNAL Free Radical Biol. Med. 27, 193-202 (1999)
AUTHORS
TITLE Alterations of SAG mRNA in human cancer cell lines: requirement for
the RING finger domain for apoptosis protection
JOURNAL Carcinogenesis. 20 (10), 1899-1903 (1999)
MEDLINE 99435944

None

AF092878 Homo sapi
AF164679 Homo sapi
BC008627 Homo sapi
BC005966 Homo sapi
AC108040 Homo sapi
AC092418 Homo sapi
AC104983 Homo sapi
AC068693 Homo sapi
AC092978 Homo sapi
AC067952 Homo sapi
AC060766 Homo sapi
AF142060 Homo sapi
AX282592 Sequence
AF092877 Mus muscu
BC011127 Mus muscu
AX192664 Sequence
AX192658 Sequence
AF312226 Homo sapi
AF323208 Homo sapi
AC067952 Homo sapi
AC106176 Rattus no
AC098496 Rattus no
AC068693 Homo sapi
AC097192 Rattus no
AC020266 Drosophill
AC011696 Drosophill
AC007473 Drosophill
AE003825 Drosophill
AC020855 Mus muscu
AX281690 Sequence
BC017370 Homo sapi
AF142059 Homo sapi
AF085906 Homo sapi
AF140598 Homo sapi
BC001466 Homo sapi
AF140599 Mus muscu
X73608 H.sapiens m
AY070810 Drosophill
AF218290 Drosophill
AY061302 Drosophill
AC019742 Drosophill
AC010010 Drosophill
AE003468 Drosophill
AX212267 Sequence
U80449 Caenorhabdit

PUBMED 10506102
 REFERENCE 3 (bases 1 to 754)
 AUTHORS Duan,H., Wang,Y., Avlram,M., Swaroop,M., Loo,J.A., Bian,J., Tian,Y., Mueller,T., Bisgaier,C.L. and Sun,Y.
 TITLE SAG, a novel zinc RING finger protein that protects cells from apoptosis induced by redox agents
 JOURNAL Molecular and cellular biology. 19 (4), 3145-3155 (1999)
 MEDLINE 99182502
 PUBMED 10082581
 REFERENCE 4 (bases 1 to 754)
 AUTHORS Swaroop,M., Wang,Y., Miller,P., Duan,H., Jatkoe,T., Madore,S.J. and Sun,Y.
 TITLE Yeast homolog of human SAG/ROC2/Rbx2/Hrt2 is essential for cell growth, but not for germination: chip profiling implicates its role in cell cycle regulation
 JOURNAL Oncogene. 19 (24), 2855-2866 (2000)
 MEDLINE 20309864
 PUBMED 10851089
 REFERENCE 5 (bases 1 to 754)
 AUTHORS Duan,H., Tsvetkov,L.M., Liu,Y., Song,Y., Swaroop,M., Wen,R., Kung,H.P., Zhang,H. and Sun,Y.
 TITLE Promotion of S-phase entry and cell growth under serum starvation by SAG/ROC2/Rbx2/Hrt2, an E3 ubiquitin ligase component: association with inhibition of p27 accumulation
 JOURNAL Molecular carcinogenesis. 30 (1), 37-46 (2001)
 MEDLINE 21152847
 PUBMED 11255262
 REFERENCE 6 (bases 1 to 754)
 AUTHORS Sun,Y.
 TITLE Direct Submission
 JOURNAL Submitted (16-SEP-1998) Department of Molecular Biology, Parke-Davis, 2800 Plymouth Rd, Ann Arbor, MI 48105, USA
 FEATURES
 source
 1..754
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="3"
 /map="3q22-q24"
 /cell_line="Hela, D98/AH-2, HPRt-"
 1..342
 /function="growth promotion"
 /note="redox sensitive, metal binding; expression protects cells from apoptosis induced by redox compounds"
 /codon_start=1
 /product="zinc RING finger protein SAG"
 /protein_id="AAB25962.1"
 /db_xref="GI:4588034"
 /translation="MADVEDGEETCALASHSGSSGSGDKMFSLKKNNAVAMNSWDVECDTCAICRVQMDACLKQAEKNQEDCVVWVGECNHSFHNCCMSLWVKNNRCPLCQDDWVVRIGK"
 BASE COUNT 205 a 155 c 201 g 193 t
 ORIGIN
 Query Match 100.0%; Score 754; DB 9; Length 754;
 Best Local Similarity 100.0%; Pred. No. 2.3e-202;
 Matches 754; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 241 ATCATTCCTTCCACAACCTGCTGATGTCCTGGGGTGAACAGAACATTCGTCGCT 300
 |||||||
 Db 241 AATCATTCCTTCCACAACCTGCTGATGTCCTGGGGTGAACAGAACATTCGTCGCT 300
 QY 301 CTCTGCCAGCAGGACTGGGTGGTCCAAAGATGCGCAATAGAGTGTAGAGGCTTC 360
 |||||||
 Db 301 CTCTGCCAGCAGGACTGGGTGGTCCAAAGATGCGCAATAGAGTGTAGAGGCTTC 360
 QY 361 TTAGCGCATGTTTCAGAGCCCTGGTGATCTGTATCCAGTCCCTACAAAGGCTAGA 420
 |||||||
 Db 361 TTAGCGCATGTTTCAGAGCCCTGGTGATCTGTATCCAGTCCCTACAAAGGCTAGA 420
 QY 421 ACACACAGGAGGATGATTTCTCAATAGAGGCGATGATCTGTGCTTTGGACTCAT 480
 |||||||
 Db 421 ACACACAGGAGGATGATTTCTCAATAGAGGCGATGATCTGTGCTTTGGACTCAT 480
 QY 481 CAAAGCCTTGTTAGCATTTTGTGAGTTTATCTTCAGAAATCTGTGATTAAGAGAT 540
 |||||||
 Db 481 CAAAGCCTTGTTAGCATTTTGTGAGTTTATCTTCAGAAATCTGTGATTAAGAGAT 540
 QY 541 AATTATTAAGGTGGTCCCTTCTACCTCTGTGTGTGTGTCGCGACACAGCTTAGAG 600
 |||||||
 Db 541 AATTATTAAGGTGGTCCCTTCTACCTCTGTGTGTGTGTCGCGACACAGCTTAGAG 600
 QY 601 TGCTATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
 |||||||
 Db 601 TGCTATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
 QY 661 AACAGGACGTGGAAGACAGCTTTCAGACTTTTGGATGCTTATGTTGATGATTAAGAA 720
 |||||||
 Db 661 AACAGGACGTGGAAGACAGCTTTCAGACTTTTGGATGCTTATGTTGATGATTAAGAA 720
 QY 721 GAATGTACAGTAACAATAATTAAGTGCAGTTTAA 754
 |||||||
 Db 721 GAATGTACAGTAACAATAATTAAGTGCAGTTTAA 754
 RESULT 2
 AF164679 836 bp mRNA linear PRI 14-OCT-1999
 LOCUS AF164679
 DEFINITION Homo sapiens ring finger protein CKBP1 mRNA, complete cds.
 ACCESSION AF164679
 VERSION AF164679.1 GI:5917673
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 836)
 AUTHORS Son,M.-Y., Park,J.W., Kim,Y.-S., Kang,S.-W., Marshak,D.R., Park,W. and Bae,Y.-S.
 TITLE Protein kinase CKII interacts with and phosphorylates the SAG
 JOURNAL Biochem. Biophys. Res. Commun. 263 (3), 743-748 (1999)
 MEDLINE 99443734
 REFERENCE 2 (bases 1 to 836)
 AUTHORS Son,M.-Y., Park,J.W., Kim,Y.-S., Kang,S.-W., Marshak,D.R., Park,W. and Bae,Y.-S.
 TITLE Direct Submission
 JOURNAL Submitted (01-JUL-1999) Department of Biochemistry, Kyungpook National University, Taegu 702-701, Korea
 FEATURES
 source
 1..836
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 54..395
 /note="SAG; contains an H2 ring finger motif"
 /codon_start=1
 /product="ring finger protein CKBP1"
 /protein_id="AAB53984.1"
 /db_xref="GI:5917674"
 /translation="MADVEDGEETCALASHSGSSGSGDKMFSLKKNNAVAMNSWD

VECDTCALCRVQVMDACLRCAENKQEDCVVWVGECNHSFHNCKMSLWVKONNRCLC
 ODDWVVRIGK*
 BASE COUNT 228 a 185 c 223 g 200 t
 ORIGIN

Query Match 91.6%; Score 691; DB 9; Length 816;
 Best Local Similarity 97.6%; Pred. No. 1.5e-184;
 Matches 744; Conservative 0; Mismatches 10; Indels 7; Gaps 4;

1 ATGGCCGAGCTGGAGAGCAGAGAAACCTGCGCCCTGCTCCTCCTCGGAGACTCA 60
 |||||||
 54 ATGGCCGAGCTGGAGAGCAGAGAAACCTGCGCCCTGCTCCTCCTCGGAGACTCA 113
 |||||||
 61 GGCTCCAACTCGGAGAGCAGAGAAATGCTCTCCCTCAAGAAGTGGAAACCGGCGCATG 120
 |||||||
 114 GGCTCCAACTCGGAGAGCAGAGAAATGCTCTCCCTCAAGAAGTGGAAACCGGCGCATG 173
 |||||||
 121 TGGAGCTGGAGAGCTGGAGATGCGATACGTCGCCATCTGCAGAGGTCCAGGTGATGATGCC 180
 |||||||
 174 TGGAGCTGGAGAGCTGGAGATGCGATACGTCGCCATCTGCAGAGGTCCAGGTGATGATGCC 233
 |||||||
 181 TGTCTTAGATGTCACAGCTCAAAACAAAGAGAGACTGTGTGTGTGTGTGTGTGTGTGTGT 240
 |||||||
 234 TGTCTTAGATGTCACAGCTCAAAACAAAGAGAGACTGTGTGTGTGTGTGTGTGTGTGTGT 293
 |||||||
 241 AATCATTTCTTCACAACTGCTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
 |||||||
 294 AATCATTTCTTCACAACTGCTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 353
 |||||||
 301 CTCCTGCGACAGAGACTGGGTGTGTCAAAGAAATGCGCAATGAGAGTGTGTAGAGGCTTC 360
 |||||||
 354 CTCCTGCGACAGAGACTGGGTGTGTCAAAGAAATGCGCAATGAGAGTGTGTAGAGGCTTC 413
 |||||||
 361 TTAGGCACTTTGTTCAGAGCCCTGCTGATCTTTGATTCAGAGTCCCTTCAAGAGCTTACA 420
 |||||||
 414 TTAGGCACTTTGTTCAGAGCCCTGCTGATCTTTGATTCAGAGTCCCTTCAAGAGCTTACA 473
 |||||||
 421 ACACCTACAGGAGTGAATCTTCAAAATAGAGCCGATGATCTGTGTC--TTTGGAGTC 478
 |||||||
 474 ACACCTACAGGAGTGAATCTTCAAAATAGAGCCGATGATCTGTGTC--TTTGGAGTC 533
 |||||||
 479 ATCAAAGCCTTGTT--ACATTTGTCACTTTATCTTCAGAAATCTCTGTGATTAAGA 536
 |||||||
 534 ATCAAAGCCTTGTTAGCATTTTGTCACTTTATCTTCAGAAATCTCTGTGATTAAGA 593
 |||||||
 537 AATATATTTATTAAGAGTGTCT 596
 |||||||
 594 AATATATTTATTAAGAGTGTCT 653
 |||||||
 597 GAAGTGTCTATTAAGAAAGAGAGCTCCAAATTTGAATCAGC--TTATATTTACCCATTTC 655
 |||||||
 654 GAAGTGTCTATTAAGAAAGAGAGAGCTCCAAATTTGAATCAGC--TTATATTTACCCATTTC 713
 |||||||
 656 TATACACAGCAGTGGAGAGCAGTTTC--GAGACTTTTTCGATGCTTATGTTGATGACT 713
 |||||||
 714 TATACACAGCAGTGGAGAGCAGTTTCGATGCTTATGTTGATGACT 773
 |||||||
 714 TAAAAAGAGTGTACAGTAAACAAATTAAGTCACTTTAAA 754
 |||||||
 774 TAAAAAGAGTGTACAGTAAACAAATTAAGTCACTTTAAA 814
 |||||||

RESULT 3
 BC008627 816 bp mRNA linear PRI 12-JUL-2001
 LOCUS
 DEFINITION Homo sapiens, ring finger protein 7, clone MGC:117274 IMAGE:4177613,
 mRNA, complete cds.
 ACCESSION BC008627
 VERSION BC008627.1 GI:14250388
 KEYWORDS MGC.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE

1 (bases 1 to 816)
 Strausberg, R.
 Direct Submission
 Submitted (25-MAY-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

REMARK

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: David N. Louis, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: villalobos@bcm.tmc.edu
 Villalobos, D.K., Luna, R.A., Hale, S.M., Huylk, S., Lu, X., Garcia,
 A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,
 Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LINL at: <http://image.llnl.gov>
 Series: IRK Plate: 12 Row: h Column: 14
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Similarity but not
 identity to protein.

FEATURES

Location/Qualifiers
 1..816
 /organism="Homo sapiens"
 /db_xref="LOCUSID:9616"
 /db_xref="taxon:9606"
 /clone="MGC:117274 IMAGE:4177613"
 /tissue_type="Brain, anaplastic oligodendroglioma with
 1p/19q loss"
 /clone_lib="NCI CGAP_Brn67"
 /lab_host="DH10B"
 /note="vector: PCMV-SPORT6"
 37..378
 /codon_start=1
 /product="ring finger protein 7"
 /protein_id="FAH08627.1"
 /db_xref="GI:14250389"
 /translation="MADVEDERTCALASHSGSSGSKSGDKMFSLKRMNAVAMMSMD
 VECDTCALCRVQVMDACLRCAENKQEDCVVWVGECNHSFHNCKMSLWVKONNRCLC
 ODDWVVRIGK"
 ODDWVVRIGK

CDS

BASE COUNT 226 a 178 c 213 g 199 t
 ORIGIN

Query Match 91.4%; Score 689.4; DB 9; Length 816;
 Best Local Similarity 97.6%; Pred. No. 4.5e-184;
 Matches 743; Conservative 0; Mismatches 11; Indels 7; Gaps 4;

1 ATGGCCGAGCTGGAGAGCAGAGAAACCTGCGCCCTGCTCCTCCTCGGAGACTCA 60
 |||||||
 37 ATGGCCGAGCTGGAGAGCAGAGAAACCTGCGCCCTGCTCCTCCTCGGAGACTCA 96
 |||||||
 61 GGCTCCAACTCGGAGAGCAGAGAAATGCTCTCCCTCAAGAAGTGGAAACCGGCGCATG 120
 |||||||
 97 GGCTCCAACTCGGAGAGCAGAGAAATGCTCTCCCTCAAGAAGTGGAAACCGGCGCATG 156
 |||||||
 121 TGGAGCTGGAGAGCTGGAGATACGTCGCCATCTGCAGAGGTCCAGGTGATGATGCC 180
 |||||||
 157 TGGAGCTGGAGAGCTGGAGATACGTCGCCATCTGCAGAGGTCCAGGTGATGATGCC 216
 |||||||
 181 TGTCTTAGATGTCACAGCTCAAAACAAAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 240
 |||||||
 217 TGTCTTAGATGTCACAGCTCAAAACAAAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 276
 |||||||

QY	241	AATCATCTCTCCACAACCTGCTGATGTCCTGCTGGGTGAAGAACAAATCGTGCCT	300
Db	277	AAATATCTCTCTCCACAACCTGCTGATGTCCTGCTGGGTGAAGAACAAATCGTGCCT	336
QY	301	CTCGCAGCAGGACCTGGTGGTCCAAAGATCGGCAATGAGTGGTTAAGAGGCTTC	360
Db	337	CTCTGCCAGCAGGACCTGGTGGTCCAAAGATCGGCAATGAGTGGTTAAGAGGCTTC	396
QY	361	TTAGCCGAGTTGTTCAAGAGCCTGCTGATCTTGAATCCAGTCCCTACAAAGCTAGA	420
Db	397	TTAGCCGAGTTGTTCAAGAGCCTGCTGATCTTGAATCCAGTCCCTACAAAGCTAGA	456
QY	421	ACACTACAGGGAGTGAATTTCTCAATAGAGCCGAGATGCTGTGCTC--TTTGAAGTC	478
Db	457	ACACTACAGGGAGTGAATTTCTCAATAGAGCCGAGATGCTGTGCTC--TTTGAAGTC	516
QY	479	ATCAACCCCTTGGTT--AGCATTTGTCAGTTTATCTTCGAATCTCTGATTTAGA	536
Db	517	ATCAACCCCTTGGTTAGCATTTTGTGATTTTATCTTCGAATTTCTGCGATTTAGA	576
QY	537	AGATTAATTTTAAGGTGGTCTCTCTACCTCTGTGTGTGTGTCGCGACACAGCTTA	596
Db	577	AGATTAATTTTAAGGTGGTCTCTCTACCTCTGTGTGTGTGTCGCGACACAGCTTA	636
QY	597	GAAGTGTATTAAGAAAGAAAGAGCTCCAAATGATGATCACC--TTATTAATTTACCATTTTC	655
Db	637	GAAGTGTATTAAGAAAGAAAGAGCTCCAAATGATGATCACC--TTATTAATTTACCATTTTC	696
QY	656	TATACACAGGAGCTGAGACAGATTTC--GAGACTTTTTCGATGCTTATGTTGATCAGT	713
Db	697	TATACACAGGAGCTGAGACAGATTTCGAGAACTTTTTCGATGCTTATGTTGATCAGT	756
QY	714	TAAGAAAGATGTTACAGTAAACAATAAGTGCAGTTTAA	754
Db	757	TAAGAAAGATGTTACAGTAAACAATAAGTGCAGTTTAA	797
RESULT	4		
LOCUS	BC005966	822 bp	mRNA
DEFINITION			822 bp
DEFINITION			mRNA, complete cds.
ACCESSION	BC005966		
VERSION	BC005966.1	GI:13543635	
KEYWORDS			MGC.
SOURCE			human.
ORGANISM			Homo sapiens
REFERENCE			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS			Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE			1 (bases 1 to 822)
JOURNAL			Strausberg, R.
			Direct Submission
			Submitted (02-APR-2001) National Institutes of Health, Mammalian
			Gene Collection (MGC), Cancer Genomics Office, National Cancer
			Institute, 31 Center Drive, Room 11A05, Bethesda, MD 20892-2590,
			USA
REMARK			
COMMENT			

NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: ATCC

CDNA Library Preparation: CLONTECH Laboratories, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www.shgc.stanford.edu>
 Contact: (Dickson, Mark) mcdepaxl.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://lmln.gov>
 Series: IRAL Plate: 21 Row: 0 Column: 13
 This clone was selected for full length sequencing because it

		passed the following selection criteria: Genomescan gene prediction, similarity but not identity to protein.	
FEATURES		Location/Qualifiers	
source		1..822	
		/organism="Homo sapiens"	
		/db_xref="LocustID:9616"	
		/db_xref="taxon:9606"	
		/clone="MGC:14618 IMAGE:4069078"	
		/tissue_type="kidney, hypernephroma"	
		/clone_lib="NIH_MGC_58"	
		/lab_host="DH10B"	
		/note="Vector: pDNR-LIB"	
CDS		27..368	
		/product="ring finger protein 7"	
		/protein_id="AA05966.1"	
		/db_xref="GI:13543636"	
		/translation="MADVEDGETCALASHSSSGSKSGGDKKFSLKNNVAAMNSWD VECDTCARVQYMDCLRCQAEKQEDCVVWVECNHSPHNCMSLWVKNNRCPLC QDDWVVRIGK"	
BASE COUNT		240 a 175 c 210 g 197 t	
ORIGIN			
Query Match		88.0%; Score 663.6; DB 9; Length 822;	
Best Local Similarity		97.1%; Pred. No. 9e-177;	
Matches 740; Conservative		0; Mismatches 14; Indels 8; Gaps 6;	
QY	1	ATGGCCAGCTGGAGAGGAGAGAACTGGCCCTGCTCTACTCCGGAGCTCA	60
Db	27	ATGGCCAGCTGGAGAGGAGAGAACTGGCCCTGCTCTACTCCGGAGCTCA	86
QY	61	GGCTCCAGTGGAGAGGAGAGATGTTCTCCCTCAAGAGTGAAGCGGCTGCCATG	120
Db	87	GGCTCCAGTGGAGAGGAGAGATGTTCTCCCTCAAGAGTGAAGCGGCTGCCATG	146
QY	121	TGAGAGTGGAGAGTGGAGTGCATGCGCATGTCAGAGGTCCAGGTGATGATGCC	180
Db	147	TGAGAGTGGAGAGTGGAGTGCATGCGCATGTCAGAGGTCCAGGTGATGATGCC	206
QY	181	TGCTTATGATGTCAAGTGAAGAAACAAGAGAGAGTGTGTGTGTGGGAGATGT	240
Db	207	TGCTTATGATGTCAAGTGAAGAAACAAGAGAGAGTGTGTGTGTGGGAGATGT	266
QY	241	AATCATCTCTCCACAACCTGCTGATGTCCTGCTGGGTGAAGCAATCGTGCCT	300
Db	267	AATCATCTCTCTCCACAACCTGCTGATGTCCTGCTGGGTGAAGCAATCGTGCCT	326
QY	301	CTCTGCCAGCAGGACCTGGTGGTCCAAAGATGCGCAATGAGTGGTTAGAGGCTTC	360
Db	327	CTCTGCCAGCAGGACCTGGTGGTCCAAAGATGCGCAATGAGTGGTTAGAGGCTTC	386
QY	361	TTAGCCGAGTTGTTCAAGAGCCTGCTGATCTTGAATCCAGTCCCTACAAAGCTAGA	420
Db	387	TTAGCCGAGTTGTTCAAGAGCCTGCTGATCTTGAATCCAGTCCCTACAAAGCTAGA	446
QY	421	ACACTACAGGGAGTGAATTTCTCAATAGAGCCGAGATGCTGTGCTC--TTTGAAGTC	478
Db	447	ACACTACAGGGAGTGAATTTCTCAATAGAGCCGAGATGCTGTGCTC--TTTGAAGTC	506
QY	479	ATCAACCCCTTGG--TTAGCATTTGTCAGTTTATCTTCAGAAATCTCTGTGATTAAGAA	537
Db	507	ATCAACCCCTTGGTTTAGCATTTTGTGATTTCTTCAGAAATCTCTGTGATTAAGAA	566
QY	538	GATTAATTTATTAAGAGTGGTCTCTCTACTCTGTGTGTGTGTGTGTGTGTGTGTGT	597
Db	567	GATTAATTTATTAAGAGTGGTCTCTCTACTCTGTGTGTGTGTGTGTGTGTGTGTGT	626
QY	598	AAATGCTAT--AAAGAGAGAGAGCTCC--AAATGATCACC--TTAATTTTACCATTT	654
Db	627	AAATGCTATTAAGAGAGAGAGCTCCAAATTTGATCACCCTTTAATTTTACCATTT	686
QY	655	CTATACACAGGAGTGGAGAGAGTTC--GAGACTTTTTCGATGCTTATGTTGATCAG	712

Db 687 CTATACACAGCAGTGGAGCAGTTTCAGAGACCTTTTGCATGCTATGTTGATCAG 746
QY 713 TTTAAAAAGATGTTACATACAAATTAAGTCAGTTAAA 754
Db 747 TTTAAAAAGATGTTACATACAAATTAAGTCAGTTAAA 788

RESULT 5
AC108040/c 163521 bp DNA linear HTG 30-JAN-2002
LOCUS Homo sapiens chromosome 4 clone RP11-210010, WORKING DRAFT
DEFINITION SEQUENCE, 4 unordered pieces.
ACCESSION AC108040
VERSION AC108040.2 GI:18425316
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 163521)
Waterston, R.H.
The sequence of Homo sapiens clone
2 (bases 1 to 163521)
Waterston, R.H.
Unpublished
Direct Submission
Submitted (24-JAN-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Jan 30, 2002 this sequence version replaced gi:18308830.

COMMENT
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@watson.wustl.edu
Project Information
Center project name: H.NH0210010

Summary Statistics

Sequencing vector: M13; 0%
Sequencing vector: plasmid; 100%
Chemistry: Dye-primer ET; 0% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 162304 bases at least Q40
Consensus quality: 162444 bases at least Q30
Consensus quality: 162529 bases at least Q20
Insert size: 189000; agarose-fp
Insert size: 163221; sum-of-ctrls
Quality coverage: 8.88 in Q20 bases; agarose-fp
Quality coverage: 8.85 in Q20 bases; sum-of-ctrls

* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1564: contig of 1564 bp in length
* 1665 1664: gap of unknown length
* 1665 88984: contig of 87320 bp in length
* 88985 89084: gap of unknown length
* 89085 112702: contig of 23618 bp in length
* 112703 112802: gap of unknown length
* 112803 163521: contig of 50719 bp in length.
Location/Qualifiers
1..163521
/organism="Homo sapiens"
/db_xref="taxon:9606"

FEATURES
SOURCE

/chromosome="4"
/clone="RP11-210010"
1..1564
/note="assembly_name:Contig25
clone_end:SP6
vector_side:left"
1665..88984
/note="assembly_name:Contig28
clone_end:T7
vector_side:left"
89085..112702
/note="assembly_name:Contig26"
112803..163521
/note="assembly_name:Contig27"
BASE COUNT 51049 a 32426 c 33346 g 46399 t 301 others
ORIGIN

Query Match 80.5%; Score 607.2; DB 2; Length 163521;
Best Local Similarity 90.6%; Pred. No. 3.2e-160;
Matches 696; Conservative 0; Mismatches 58; Indels 14; Gaps 4;

QY 1 ATGGCCGACGTGGAGAGCAGAGAAACCTGGCCCTGCTCCTCCTCGGAGCTCA 60
Db 80418 ATGGCCGACGTGGAGAGCAGAGAAACCTGACCTGACCTCCTCCTCGGAGCTCA 80359
QY 61 GGCTCCAGTCCGGAGGCGCAAAAGATGTCCTCCCAAGAGTGAACCGGTCGATG 120
Db 80358 GGCTCCAGTCCGGAGGCGCAAAAGATGTCCTCCCAAGAGTGAACCGGTCGATG 80299
QY 121 TGGAGCTGGAGCTGGAGTGGCATTACGTCGCGCATTCGAGGTCGAGTGCATGCC 180
Db 80298 TGGAGCTGGAGCTGGAGTGGCATTACGTCGCGCATTCGAGGTCGAGTGCATGCC 80239
QY 181 TGTCTTATGATGTCACGCTGAAACAAAGAGAGCTGTGTGTGCTGGGAGATGT 240
Db 80238 TGTCTTATGATGTCACGCTGAAACAAAGAGAGCTGTGTGTGCTGGGAGATGT 80179
QY 241 AATCATTCCTTCACAAAGCTGTCATGTCCTGTCGTCGTCGTCGTCGTCGTCGTC 300
Db 80178 AATCATTCCTTCACAAAGCTGTCATGTCCTGTCGTCGTCGTCGTCGTCGTCGTC 80119
QY 301 CTCCTGCACGACGACGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 360
Db 80118 CTCCTGCACGACGACGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 80059
QY 361 TTAGCGCAGTTGTTCAAGCCCTGCTGATCTTGTATCCAGTCCCTCAAAAGCTTAA 420
Db 80058 TTAGCGCAGTTGTTCAAGCCCTGCTGATCTTGTATCCAGTCCCTCAAAAGCTTAA 79999
QY 421 ACACCTACAGGGGATGATCTTCGAATAGAGCGGATGATGTCGTGTC--TTTGGACTC 478
Db 79998 ACACCTACAGGGGATGATCTTCGAATAGAGCGGATGATGTCGTGTC--TTTGGACTC 79939
QY 479 ATCAAAAGCCTTGGTT--AGCATTTGTCAGTTTATCTTCAAGAAATCTCTGTGATTAAGA 536
Db 79938 ATCAAAAGCCTTGGTTAGCATTTTGTCAAGTTTATCTTCAAGAAATCTCTGATTAAGA 79879
QY 537 AGATTAATTTTAAAGGTGCTCTTCTACCTCTGTGTGTGTGTGTGTGTGTGTGTGTGT 596
Db 79878 AGATTAATTTTAAAGGTGCTCTTCTACCTCTGTGTGTGTGTGTGTGTGTGTGTGT 79819
QY 597 GAAGTCTTAAAGAAAGAAAGAGAGCCTCAATATGATACAC--TTATATTTTACCATTC 655
Db 79818 GAAGTCTTAAAGAAAG 79759
QY 656 TATACACAGCAGTGGAGAGAGCTTC-----GAGACTTTTTCAGTCTTATGATGTT 706
Db 79758 TATACACAGCAGTGGAGAGAGAGCTTCATTCAGAGAGAGAGAGAGAGAGAGAGAGAG 79699
QY 707 GATCAGTTAAAGAAATGTTACAGTTACAAATTAAGTCAGTTAAA 754
Db 79698 GATCAGTTAAAGAAATGTTACAGTTACAAATTAAGTCAGTTAAA 79651

```

RESULT      6
LOCUS       AC092418/c
DEFINITION  Homo sapiens chromosome 3 clone RP11-229A12, WORKING DRAFT
ACCESSION   AC092418
VERSION     AC092418.1
KEYWORDS    HTGS_PHADEL; HTGS_DRAFT; HTGS_FULLTOP; HTGS_ACTIVEFIN.
SOURCE      human.
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 163757)
AUTHORS     Kaul, R.K., Olson, M.V., Raymond, C., Clendinning, J., Ivey, R.G. and
             Haugen, E.D.
TITLE       Direct Submission
JOURNAL     Unpublished
AUTHORS     2 (bases 1 to 163757)
            Kaul, R.K., Olson, M.V., Raymond, C., Clendinning, J., Ivey, R.G. and
            Haugen, E.D.
TITLE       Direct Submission
JOURNAL     Submitted (04-JUL-2001) Genome Center, University of Washington,
            Box 352145, Seattle, WA 98195, USA
            On Jul 4, 2001 this sequence version replaced g1:8705848.
COMMENT     -----
            Center: University of Washington Genome Center
            Web site: http://www.genome.washington.edu
            Contact: uwgchgs@u.washington.edu
            Drafting Center: BCM
            -----
            Project Information
            Center project name: chr-3
            Center clone name: RP11-229A12 (bc0315)
            -----
            Summary Statistics
            Sequencing vector: plasmid; 50% of reads
            Chemistry: Dye-terminator ET; 47% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 162824 bases at least Q40
            Consensus quality: 163281 bases at least Q30
            Consensus quality: 163468 bases at least Q20
            Insert size: 16357, Sum-of-contigs
            Quality coverage: 7.6x in Q20 bases; sum-of-contigs
            -----
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 3 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *
            * 1 15071: contig of 15071 bp in length
            * 15072 15171: gap of unknown length
            * 15172 57866: contig of 42715 bp in length
            * 57867 57986: gap of unknown length
            * 57987 163757: contig of 105771 bp in length.
            Location/Qualifiers
            1.163757
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="3"
            /clone="RP11-229A12"
            /clone_lib="RP11 human BAC library 11"
            1.15071
            /note="assembly_name:Contig6"
            15172..57866
            /note="assembly_name:Contig7"
            57987..163757
            /note="assembly_name:Contig8"
            -----
FEATURES
source
1.163757
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="3"
/clone="RP11-229A12"
/clone_lib="RP11 human BAC library 11"
1.15071
/note="assembly_name:Contig6"
15172..57866
/note="assembly_name:Contig7"
57987..163757
/note="assembly_name:Contig8"

```


* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved. 724: contig of 724 bp in length
1
725 824: gap of 100 bp
825 1582: contig of 758 bp in length
1583 1682: gap of 100 bp
1683 2409: contig of 727 bp in length
2410 2509: gap of 100 bp
2510 3235: contig of 726 bp in length
3236 3335: gap of 100 bp
3336 4091: contig of 755 bp in length
4091 4190: gap of 100 bp
4191 4955: contig of 765 bp in length
4956 5055: gap of 100 bp
5056 5780: contig of 725 bp in length
5781 5880: gap of 100 bp
5881 6643: contig of 763 bp in length
6644 6743: gap of 100 bp
6744 7496: contig of 753 bp in length
7497 7596: gap of 100 bp
7597 8348: contig of 752 bp in length
8349 8448: gap of 100 bp
8449 9217: contig of 769 bp in length
9218 9317: gap of 100 bp
9318 10093: contig of 776 bp in length
10094 10193: gap of 100 bp
10194 10924: contig of 731 bp in length
10925 11024: gap of 100 bp
11025 11793: contig of 769 bp in length
11794 11893: gap of 100 bp
11894 12661: contig of 768 bp in length
12662 12761: gap of 100 bp
12762 13525: contig of 764 bp in length
13526 13625: gap of 100 bp
13626 14347: contig of 722 bp in length
14348 14447: gap of 100 bp
14448 15172: contig of 725 bp in length
15173 15272: gap of 100 bp
15273 16007: contig of 735 bp in length
16008 16107: gap of 100 bp
16108 16855: contig of 748 bp in length
16856 16955: gap of 100 bp
16956 17711: contig of 756 bp in length
17712 17811: gap of 100 bp
17812 18568: contig of 757 bp in length
18569 18668: gap of 100 bp
18669 19421: contig of 753 bp in length
19422 19521: gap of 100 bp
19522 20229: contig of 708 bp in length
20230 20329: gap of 100 bp
20330 21061: contig of 732 bp in length
21062 21161: gap of 100 bp
21162 21909: contig of 748 bp in length
21910 22009: gap of 100 bp
22010 22773: contig of 764 bp in length
22774 22873: gap of 100 bp
22874 23623: contig of 750 bp in length
23624 23723: gap of 100 bp
23724 24489: contig of 766 bp in length
24490 24589: gap of 100 bp
24590 25354: contig of 765 bp in length
25355 25454: gap of 100 bp
25455 26174: contig of 720 bp in length
26175 26274: gap of 100 bp
26275 27043: contig of 769 bp in length
27044 27143: gap of 100 bp
27144 27900: contig of 757 bp in length
27901 28000: gap of 100 bp
28001 28753: contig of 753 bp in length
28754 28853: gap of 100 bp
28854 29572: contig of 719 bp in length

* 29573 29672: gap of 100 bp
29673 30419: contig of 747 bp in length
30420 30519: gap of 100 bp
30520 31262: contig of 743 bp in length
31263 31362: gap of 100 bp
31363 32052: contig of 690 bp in length
32053 32152: gap of 100 bp
32153 32909: contig of 757 bp in length
32910 33009: gap of 100 bp
33010 33737: contig of 727 bp in length
33737 33836: gap of 100 bp
33837 34598: contig of 762 bp in length
34599 34698: gap of 100 bp
34699 35456: contig of 758 bp in length
35457 35556: gap of 100 bp
35557 36274: contig of 718 bp in length
36275 36374: gap of 100 bp
36375 37111: contig of 737 bp in length
37112 37211: gap of 100 bp
37212 37954: contig of 743 bp in length
37955 38054: gap of 100 bp
38055 38783: contig of 729 bp in length
38784 38883: gap of 100 bp
38884 39686: contig of 803 bp in length
39687 39786: gap of 100 bp
39787 40535: contig of 749 bp in length
40536 40635: gap of 100 bp
40636 41387: contig of 752 bp in length
41388 41487: gap of 100 bp
41488 42254: contig of 767 bp in length
42255 42354: gap of 100 bp
42355 43046: contig of 692 bp in length
43047 43146: gap of 100 bp
43147 43885: contig of 739 bp in length
43886 43985: gap of 100 bp
43986 44741: contig of 756 bp in length
44742 44841: gap of 100 bp
44842 45604: contig of 763 bp in length
45605 45704: gap of 100 bp
45705 46429: contig of 725 bp in length
46430 46529: gap of 100 bp
46530 47244: contig of 715 bp in length
47245 47344: gap of 100 bp
47345 48088: contig of 744 bp in length
48089 48188: gap of 100 bp
48189 48948: contig of 760 bp in length
48949 49048: gap of 100 bp
49049 49803: contig of 755 bp in length
49804 49903: gap of 100 bp
49904 50648: contig of 745 bp in length
50649 50748: gap of 100 bp
50749 51506: contig of 758 bp in length
51507 51606: gap of 100 bp
51607 52345: contig of 739 bp in length
52346 52445: gap of 100 bp
52446 53185: contig of 740 bp in length
53186 53285: gap of 100 bp
53286 54059: contig of 774 bp in length
54060 54159: gap of 100 bp
54160 54901: contig of 742 bp in length
54902 55001: gap of 100 bp
55002 55740: contig of 739 bp in length
55741 55840: gap of 100 bp
55841 56577: contig of 737 bp in length
56578 56677: gap of 100 bp
56678 57433: contig of 756 bp in length
57434 57533: gap of 100 bp
57534 58290: contig of 757 bp in length

Query Match 73.4% Score 553.8: DB 2: Length 59155;
Best Local Similarity 89.9% Pred. No. 3.2e-145;
Matches 642; Conservative 0; Mismatches 58; Indels 14; Gaps 4;

```

OY 45 TCACGCCGAGAGCTCAGGCTCCAGAGTGGAGGCGCAAGATGTTCTCCCTCAAGAGT 104
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 7635 TCACCTCAGAGAGGCGAGGCTCCAGTGGAGGCGCAAGATGTTCTTACTCAAGAGT 7694
OY 105 GAACCGCGTCCAGTGTGAGCTGGAGCTGGAGTGGATACGTGGCCATCTGCAGGCT 164
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 7695 GAACCGTGGTGGAGTGTGAGCTGGAGCTGGAGTGGATACGTGGCCATCTGCAGGCT 7754
OY 165 CCAGGTGAGTGGATGCTGCTTGAATGCAAGCTGAAACCAAGAGAGAGTGTCT 224
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 7755 CCAGGTGAGTGGATGCTGCTTGAATGCAAGCTGAAACCAAGAGAGAGTGTCT 7814
OY 225 GGTCTGGAGAGATGATATCATCTCTTCCCAACTGCTGATGCTGCTGGTGAACA 284
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 7815 GGTCTGGAGAGATGATATCATCTCTTCCCAACTGCTGATGCTGCTGGTGAACA 7874
OY 285 GAACAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 344
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 7875 GAACAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7934
OY 345 GTGGTGAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 404
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 7935 GTGGTGAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7994
OY 405 CCTTCAAAAGGCTAGAACACTACAGGAGATGATGCTTCAATAGAGAGGAGTGTG 464
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 7995 CCTTCAAAAGGCTAGAACACTACAGGAGATGATGCTTCAATAGAGAGGAGTGTG 8054
OY 465 TGGTC--TTTGGAGCTATCAAGAGCTTGGT--AGCATTTGTCAGTTTATCTTGAGAA 520
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 8055 TGGTCCTTTGGAGACATCAAGAGCTTGGT--AGCATTTGTCAGTTTATCTTGAGAA 8114
OY 521 TTCTCTGATTAAGAAATATTTATTAAGTGTCTCTTCTGCTGCTGCTGCTGCTG 580
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 8115 TTCTCTGATTAAGAAATATTTATTAAGTGTCTCTTCTGCTGCTGCTGCTGCTG 8174
OY 581 TGGCCACACAGCTTGAAGTGTCTTAAAGAAAGAGAGCTTCAATGATGATCACC-TT 639
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 8175 TGGCCACACAGCTTGAAGTGTCTTAAAGAAAGAGAGCTTCAATGATGATCACC-TT 8234
OY 640 ATAATTTACCAATTTCTATACAAAGAGAGTGAAGAGCTTTC-----GAGACTTT 690
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 8235 ATAATTTACCAATTTCTATACAAAGAGAGTGAAGAGCTTTC-----GAGACTTT 8294
OY 691 TTGCAGTGTATGTTGATGATCACTTAAAGAAAGATGTTACAGTAAATAAAGT 744
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 8295 TTGCAGTGTATGTTGATGATCACTTAAAGAAAGATGTTACAGTAAATAAAGT 8348

```

```

RESULT 8
AC068693/C 125041 bp DNA linear HTG 04-MAY-2001
LOCUS Homo sapiens chromosome 17 clone RP11-144C9 map 17, WORKING DRAFT
DEFINITION
SEQUENCE, 25 unordered pieces.
AC068693
AC068693.2 GI:8139716
VERSION
KEYWORDS
SOURCE
human.
ORGANISM
Homo sapiens
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 125041)
Biren, B., Linton, L., Nussbaum, C. and Lander, E.
JOURNAL
TITLE
Homo sapiens chromosome 17, clone RP11-144C9
AUTHORS
2 (bases 1 to 125041)
Biren, B., Linton, L., Nussbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,
Boudjovskiy, L., Boukhalter, B., Brown, A., Burkett, G.,
Campoliano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,
Collamore, A., Cooke, P., DeArrelano, K., Dewar, K., Diaz, J.S.,
Dodge, S., Domingo, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,
Galagan, J., Gardyna, S., Glade, S., Goyette, M., Graham, L.,

```

TITLE JOURNAL COMMENT

```

Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howard, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
Klein, J., Labrecque, K., Lamazares, R., Landers, T., Lehotzky, J.,
Levine, R., Liu, C., Liu, G., Locke, K., MacDonald, P., Margulis, N.,
McCarthy, M., McEwan, P., McGovern, A., McKernan, K., McSheeters, R.,
Meidrum, J., Meneus, L., Mihova, T., Miranda, C., Mienna, V., Morrow, J.,
Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N.,
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Testafaye, S., Theodore, J., Tirrell, A., Travers, M., Triggillo, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
Young, G., Zainoun, J., Zimmer, A. and Zody, M.
Direct Submission
Submitted (06-MAY-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 1, 2000 this sequence version replaced gi:171215.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/MW/repeatmasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Web site: http://www-seq.wi.mit.edu
Center code: WIBR
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: 144_C-9
Center clone name: 18985
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 112098 bases at least Q40
Consensus quality: 118849 bases at least Q30
Consensus quality: 121119 bases at least Q20
Insert size: 158000; agarose-efp
Insert size: 122641; sum-of-coverage
Quality coverage: 3.8 in Q20 bases; agarose-efp
Quality coverage: 4.9 in Q20 bases; sum-of-coverage
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 25 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 1019: contig of 1019 bp in length
* 1020 1119: gap of 100 bp
* 1120 3013: contig of 1894 bp in length
* 3014 3113: gap of 100 bp
* 3114 5008: contig of 1895 bp in length
* 5009 5108: gap of 100 bp
* 5109 7154: contig of 2046 bp in length
* 7155 7254: gap of 100 bp
* 7255 10405: contig of 3151 bp in length
* 10406 10505: gap of 100 bp
* 10506 12914: contig of 2409 bp in length
* 12915 13014: gap of 100 bp
* 13015 13781: contig of 767 bp in length
* 13782 13881: gap of 100 bp
* 13882 17746: contig of 3865 bp in length
* 17747 17846: gap of 100 bp
* 17847 20494: contig of 2648 bp in length
* 20495 22173: gap of 100 bp
* 22174 22273: contig of 1579 bp in length
* 22274 25788: contig of 3515 bp in length
* 25789 25888: gap of 100 bp
* 25889 30218: contig of 4330 bp in length
* 30219 30318: gap of 100 bp

```

```

* 30319 33343: contig of 3025 bp in length
* 33344 33443: gap of 100 bp
* 33444 36428: contig of 2985 bp in length
* 36429 36528: gap of 100 bp
* 36529 40028: contig of 3500 bp in length
* 40029 40128: gap of 100 bp
* 40129 45543: contig of 5415 bp in length
* 45544 45643: gap of 100 bp
* 45644 52370: contig of 6727 bp in length
* 52371 52470: gap of 100 bp
* 52471 56869: contig of 4399 bp in length
* 56870 56968: gap of 100 bp
* 56970 62452: contig of 5483 bp in length
* 62453 62552: gap of 100 bp
* 62553 70768: contig of 8216 bp in length
* 70769 70868: gap of 100 bp
* 70869 76175: contig of 5307 bp in length
* 76176 76275: gap of 100 bp
* 76276 84628: contig of 8353 bp in length
* 84629 84728: gap of 100 bp
* 84729 95400: contig of 10672 bp in length
* 95401 95500: gap of 100 bp
* 95501 111561: contig of 16061 bp in length
* 111562 111661: gap of 100 bp
* 111662 125041: contig of 13380 bp in length.

```

FEATURES

SOURCE

```

1..125041
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /chromosome="17"
  /map="17"
  /clone="RP11-14AC9"
  /clone_1db="RP11 Human Male BAC"
  1..1019
    /note="assembly-fragment"
    1120..3013
      /note="assembly-fragment"
      3114..5008
        /note="assembly-fragment"
        5109..7154
          /note="assembly-fragment"
          7255..10405
            /note="assembly-fragment"
            10506..12914
              /note="assembly-fragment"
              13015..13781
                /note="assembly-fragment"
                clone_end:SP6
                vector_side:right"
                13882..17746
                  /note="assembly-fragment"
                  17847..20494
                    /note="assembly-fragment"
                    20595..22173
                      /note="assembly-fragment"
                      clone_end:T7
                      vector_side:right"
                      22274..25788
                        /note="assembly-fragment"
                        25889..30218
                          /note="assembly-fragment"
                          30319..33343
                            /note="assembly-fragment"
                            33444..36428
                              /note="assembly-fragment"
                              36529..40028
                                /note="assembly-fragment"
                                40129..45543
                                  /note="assembly-fragment"
                                  45644..52370
                                    /note="assembly-fragment"
                                    52471..56869
                                      /note="assembly-fragment"

```

```

misc_feature 56970..62452
              /note="assembly-fragment"
misc_feature 62553..70768
              /note="assembly-fragment"
misc_feature 70869..76175
              /note="assembly-fragment"
misc_feature 76276..84628
              /note="assembly-fragment"
misc_feature 84729..95400
              /note="assembly-fragment"
misc_feature 95501..111561
              /note="assembly-fragment"
misc_feature 111662..125041
              /note="assembly-fragment"
BASE COUNT 34737 a 26838 c 26914 g 34145 t 2407 others
ORIGIN

```

```

Query Match 62.0%; Score 467.4; DB 2; Length 125041;
Best Local Similarity 96.7%; Pred. No.1,1e-120;
Matches 521; Conservative 0; Mismatches 11; Indels 7; Gaps 4;

```

```

QY 223 GTGCTGTGGGAGAAATGTAATCATTCCTCCACACACTGCTGCATGCTGCTGGTGAA 282
DB 16007 GTGCTGTGGGAGAAATGTAATCATTCCTCCACACACTGCTGCATGCTGCTGGTGAA 15948
QY 283 CAGACATGCGTGCCTCTCCGACGAGACTGGGTGTCGAAGATCGCAATGA 342
DB 15947 CAGACATGCGTGCCTCTCCGACGAGACTGGGTGTCGAAGATCGCAATGA 15888
QY 343 GAGTGTGTAAGAGCTTCTTAGCGCAGTTGTCAGAGCCCTGGTGATCTGTAATCCAG 402
DB 15887 GAGTGTGTAAGAGCTTCTTAGCGCAGTTGTCAGAGCCCTGGTGATCTGTAATCCAG 15828
QY 403 TGGCCCTACAAAGGCTGAACTACACAGGAGATGTAATCTTCAAAATGAGCCGATGATC 462
DB 15827 TGGCCCTACAAAGGCTGAACTACACAGGAGATGTAATCTTCAAAATGAGCCGATGATC 15768
QY 463 TGTGCTC--TTTGACTCATCAAGCCTGGTTAGCA--TTTGCTAGTTTATCTTCAGA 518
DB 15767 TGTGCTCCTTTGGGACTCATCAAGCCTGGTTAGCA--TTTGCTAGTTTATCTTCAGA 15708
QY 519 AATTCCTGTGATTTAAGAGATATTTAAGGTGCTCTTCTTACCTCTGTGTGTG 578
DB 15707 AATTCCTGTGATTTAAGAGATATTTAAGGTGCTCTTCTTACCTCTGTGTGTG 15648
QY 579 TGTGCGGACACAGCTTGAAGTGTATRAAAAAGAGAGAGCTCAATGATGATCACC- 637
DB 15647 TGTGCGGACACAGCTTGAAGTGTATRAAAAAGAGAGAGCTCAATGATGATCACC- 15588
QY 638 TTATATTTTACCATTTTCTATCAACAGCAGTGAAGAGTTTC--GAGACTTTTTCGA 695
DB 15587 TTATATTTTACCATTTTCTATCAACAGCAGTGAAGAGTTTC--GAGACTTTTTCGA 15528
QY 696 TCGTTATGTTGATCAGTTAAAAAGAAATGTTACAGTAACAATTAAGTCAGTTTAA 754
DB 15527 TCGTTATGTTGATCAGTTAAAAAGAAATGTTACAGTAACAATTAAGTCAGTTTAA 15469
RESULT 9
AC092978 179877 bp DNA linear HTG 09-AUG-2001
LOCUS Homo sapiens chromosome 3q clone RP11-34065, WORKING DRAFT
DEFINITION AC092978
ACCESSION AC092978
VERSION AC092978.1 GI:15136751
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 179877)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
AUTHORS

```

Alsbrooks,S.L., Amaralunge,H.C., Are,J.R., Banks,T., Barbara,J.,
 Benton,J., Bimega,K., Blankenburg,K., Bonnin,D., Bouck,J.,
 Bowle,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
 Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Cartron,T.F.,
 Carter,N., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
 Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
 Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
 Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
 Denn,A.L., Ding,Y., Dlinh,H.H., Douthwaite,K.J., Draper,H.,
 Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
 Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
 Foster,P., Frantz,P., Gablitz,A., Gao,J., Garcia,A., Garner,T.,
 Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
 Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
 Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
 Hollins,B., Homsif,F., Howard,S., Huber,J., Huliyk,S., Hume,J.,
 Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
 Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
 Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
 Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,C., Liu,W.,
 Louised,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
 Ma,J., Maheshwari,M., Kapua,P., Martin,R., Martindale,A.,
 Martinez,E., Massey,E., Mawhinney,E., McLeod,M.P., Meador,M.,
 Mel,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
 Morgan,M., Morris,S., Moser,M., Neal,D., Newtonson,J., Newton,N.,
 Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenwo,S.,
 Ogub,M., Okunnu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
 Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
 Quiles,M., Ren,Y., Rivers,M., Rojas,A., Rojoudan,H., Rolfe,M.,
 Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,K., Shooshari,N.,
 Slisom,I., Sodergren,E., Sonalke,T., Sparks,A., Stanley,H.,
 Stone,H., Sutton,A., Svatek,A., Taber,P., Tamerisa,A., Tamerisa,K.,
 Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,
 Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalob,D., Vinson,R.,
 Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
 Wacington,S., Williams,G., Williamson,A., Wleczky,R., Wooden,S.,
 Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
 Weinstein,G. and Gibbs,R.

Unpublished
 2 (bases 1 to 179877)
 Worley,K.C.

Direct Submission
 Submitted (09-AUG-2001) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu

----- Project Information
 Project name: HBDP
 Center clone name: RP11-340E6

----- Summary Statistics
 Assembly program: Phrap; version 0.990329
 Consensus quality: 179371 bases at least Q40
 Consensus quality: 185709 bases at least Q30
 Consensus quality: 188640 bases at least Q20
 Estimated insert size: 180121; sum-of-coverage estimation
 Quality coverage: 0x in Q20 bases; agarose-gel estimation
 Quality coverage: 3.7x in Q20 bases; sum-of-coverage estimation

----- NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 32 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

FEATURES	source	BASE COUNT	ORIGIN
1	11845: contig of 11845 bp in length	48842 a	36893 c 40248 g 50788 t 3106 others
11846	11945: gap of unknown length		
11945	11945: contig of 7616 bp in length		
11946	11945: gap of unknown length		
11947	11945: gap of unknown length		
11948	11945: gap of unknown length		
11949	11945: gap of unknown length		
11950	11945: gap of unknown length		
11951	11945: gap of unknown length		
11952	11945: gap of unknown length		
11953	11945: gap of unknown length		
11954	11945: gap of unknown length		
11955	11945: gap of unknown length		
11956	11945: gap of unknown length		
11957	11945: gap of unknown length		
11958	11945: gap of unknown length		
11959	11945: gap of unknown length		
11960	11945: gap of unknown length		
11961	11945: gap of unknown length		
11962	11945: gap of unknown length		
11963	11945: gap of unknown length		
11964	11945: gap of unknown length		
11965	11945: gap of unknown length		
11966	11945: gap of unknown length		
11967	11945: gap of unknown length		
11968	11945: gap of unknown length		
11969	11945: gap of unknown length		
11970	11945: gap of unknown length		
11971	11945: gap of unknown length		
11972	11945: gap of unknown length		
11973	11945: gap of unknown length		
11974	11945: gap of unknown length		
11975	11945: gap of unknown length		
11976	11945: gap of unknown length		
11977	11945: gap of unknown length		
11978	11945: gap of unknown length		
11979	11945: gap of unknown length		
11980	11945: gap of unknown length		
11981	11945: gap of unknown length		
11982	11945: gap of unknown length		
11983	11945: gap of unknown length		
11984	11945: gap of unknown length		
11985	11945: gap of unknown length		
11986	11945: gap of unknown length		
11987	11945: gap of unknown length		
11988	11945: gap of unknown length		
11989	11945: gap of unknown length		
11990	11945: gap of unknown length		
11991	11945: gap of unknown length		
11992	11945: gap of unknown length		
11993	11945: gap of unknown length		
11994	11945: gap of unknown length		
11995	11945: gap of unknown length		
11996	11945: gap of unknown length		
11997	11945: gap of unknown length		
11998	11945: gap of unknown length		
11999	11945: gap of unknown length		
12000	11945: gap of unknown length		
12001	11945: gap of unknown length		
12002	11945: gap of unknown length		
12003	11945: gap of unknown length		
12004	11945: gap of unknown length		
12005	11945: gap of unknown length		
12006	11945: gap of unknown length		
12007	11945: gap of unknown length		
12008	11945: gap of unknown length		
12009	11945: gap of unknown length		
12010	11945: gap of unknown length		
12011	11945: gap of unknown length		
12012	11945: gap of unknown length		
12013	11945: gap of unknown length		
12014	11945: gap of unknown length		
12015	11945: gap of unknown length		
12016	11945: gap of unknown length		
12017	11945: gap of unknown length		
12018	11945: gap of unknown length		
12019	11945: gap of unknown length		
12020	11945: gap of unknown length		
12021	11945: gap of unknown length		
12022	11945: gap of unknown length		
12023	11945: gap of unknown length		
12024	11945: gap of unknown length		
12025	11945: gap of unknown length		
12026	11945: gap of unknown length		
12027	11945: gap of unknown length		
12028	11945: gap of unknown length		
12029	11945: gap of unknown length		
12030	11945: gap of unknown length		
12031	11945: gap of unknown length		
12032	11945: gap of unknown length		
12033	11945: gap of unknown length		
12034	11945: gap of unknown length		
12035	11945: gap of unknown length		
12036	11945: gap of unknown length		
12037	11945: gap of unknown length		
12038	11945: gap of unknown length		
12039	11945: gap of unknown length		
12040	11945: gap of unknown length		
12041	11945: gap of unknown length		
12042	11945: gap of unknown length		
12043	11945: gap of unknown length		
12044	11945: gap of unknown length		
12045	11945: gap of unknown length		
12046	11945: gap of unknown length		
12047	11945: gap of unknown length		
12048	11945: gap of unknown length		
12049	11945: gap of unknown length		
12050	11945: gap of unknown length		
12051	11945: gap of unknown length		
12052	11945: gap of unknown length		
12053	11945: gap of unknown length		
12054	11945: gap of unknown length		
12055	11945: gap of unknown length		
12056	11945: gap of unknown length		
12057	11945: gap of unknown length		
12058	11945: gap of unknown length		
12059	11945: gap of unknown length		
12060	11945: gap of unknown length		
12061	11945: gap of unknown length		
12062	11945: gap of unknown length		
12063	11945: gap of unknown length		
12064	11945: gap of unknown length		
12065	11945: gap of unknown length		
12066	11945: gap of unknown length		
12067	11945: gap of unknown length		
12068	11945: gap of unknown length		
12069	11945: gap of unknown length		
12070	11945: gap of unknown length		
12071	11945: gap of unknown length		
12072	11945: gap of unknown length		
12073	11945: gap of unknown length		
12074	11945: gap of unknown length		
12075	11945: gap of unknown length		
12076	11945: gap of unknown length		
12077	11945: gap of unknown length		
12078	11945: gap of unknown length		
12079	11945: gap of unknown length		
12080	11945: gap of unknown length		
12081	11945: gap of unknown length		
12082	11945: gap of unknown length		
12083	11945: gap of unknown length		
12084	11945: gap of unknown length		
12085	11945: gap of unknown length		
12086	11945: gap of unknown length		
12087	11945: gap of unknown length		
12088	11945: gap of unknown length		
12089	11945: gap of unknown length		
12090	11945: gap of unknown length		
12091	11945: gap of unknown length		
12092	11945: gap of unknown length		
12093	11945: gap of unknown length		
12094	11945: gap of unknown length		
12095	11945: gap of unknown length		
12096	11945: gap of unknown length		
12097	11945: gap of unknown length		
12098	11945: gap of unknown length		
12099	11945: gap of unknown length		
12100	11945: gap of unknown length		
12101	11945: gap of unknown length		
12102	11945: gap of unknown length		
12103	11945: gap of unknown length		
12104	11945: gap of unknown length		
12105	11945: gap of unknown length		
12106	11945: gap of unknown length		
12107	11945: gap of unknown length		
12108	11945: gap of unknown length		
12109	11945: gap of unknown length		
12110	11945: gap of unknown length		
12111	11945: gap of unknown length		
12112	11945: gap of unknown length		
12113	11945: gap of unknown length		
12114	11945: gap of unknown length		
12115	11945: gap of unknown length		
12116	11945: gap of unknown length		
12117	11945: gap of unknown length		
12118	11945: gap of unknown length		
12119	11945: gap of unknown length		
12120	11945: gap of unknown length		
12121	11945: gap of unknown length		
12122	11945: gap of unknown length		
12123	11945: gap of unknown length		
12124	11945: gap of unknown length		
12125	11945: gap of unknown length		
12126	11945: gap of unknown length		
12127	11945: gap of unknown length		
12128	11945: gap of unknown length		
12129	11945: gap of unknown length		
12130	11945: gap of unknown length		
12131	11945: gap of unknown length		
12132	11945: gap of unknown length		
12133	11945: gap of unknown length		
12134	11945: gap of unknown length		
12135	11945: gap of unknown length		
12136	11945: gap of unknown length		
12137	11945: gap of unknown length		
12138	11945: gap of unknown length		
12139	11945: gap of unknown length		
12140	11945: gap of unknown length		
12141	11945: gap of unknown length		
12142	11945: gap of unknown length		
12143	11945: gap of unknown length		
12144	11945: gap of unknown length		
12145	11945: gap of unknown length		
12146	11945: gap of unknown length		
12147	11945: gap of unknown length		
12148	11945: gap of unknown length		
12149	11945: gap of unknown length		
12150	11945: gap of unknown length		
12151	11945: gap of unknown length		
12152	11945: gap of unknown length		
12153	11945: gap of unknown length		
12154	11945: gap of unknown length		
12155	11945: gap of unknown length		
12156	11945: gap of unknown length		
12157	11945: gap of unknown length		
12158	11945: gap of unknown length		
12159	11945: gap of unknown length		
12160	11945: gap of unknown length		
12161	11945: gap of unknown length		
12162	11945: gap of unknown length		
12163	11945: gap of unknown length		
12164	11945: gap of unknown length		
12165	11945: gap of unknown length		
12166	11945: gap of unknown length		
12167	11945: gap of unknown length		
12168	11945: gap of unknown length		
12169	11945: gap of unknown length		
12170	11945: gap of unknown length		
12171	11945: gap of unknown length		
12172	11945: gap of unknown length		
12173	11945: gap of unknown length		
12174	11945: gap of unknown length		
12175	11945: gap of unknown length		
12176	11945: gap of unknown length		
12177	11945: gap of unknown length		
12178	11945: gap of unknown length		
12179	11945: gap of unknown length		
12180	11945: gap of unknown length		
12181	11945: gap of unknown length		

Query Match 60.4%; Score 455.4; DB 2; Length 179877;
 Best Local Similarity 96.5%; Pred. No. 3e-117;
 Matches 520; Conservative 0; Mismatches 11; Indels 8; Gaps 5;

QY 223 GTGCTGGGGAGATGATATCATTCCTTCACAACTCTGATGTCCCTGTGGTGA 282
 DB 65845 GTGCTGGGGAGATGATATCATTCCTTCACAACTCTGATGTCCCTGTGGTGA 65904

QY 283 CAGAACATCGCTCCCTCTGTCAGAGACAGTGGGTGTCACAAAGATGGCAATGA 342
 DB 65905 CAGAACATCGCTCCCTCTGTCAGAGACAGTGGGTGTCACAAAGATGGCAATGA 65964

QY 343 GAGTGGTTAGAGGCTTCTTAGCCGAGTGTTCAGAGCCCTGGTGATCTTGAATCCAG 402
 DB 65965 GAGTGGTTAGAGGCTTCTTAGCCGAGTGTTCAGAGCCCTGGTGATCTTGAATCCAG 66024

QY 403 TGCCCTCAAGGCTAGAGCTACAGGGATGTAATTTCTCAATAGAGCCGATGATC 462
 DB 66025 TGCCCTCAAGGCTAGAGCTACAGGGATGTAATTTCTCAATAGAGCCGATGATC 66084

QY 463 TGTGGTC--TTTGACATCATCAAGCCCTGGTTAGC--ATTGTCAGTTTATCTTACA 518
 DB 66085 TGTGGTCCTTTGGGACATCATCAAGCCCTGGTTAGCATTGTCAGTTTATCTTACA 66144

QY 519 AATTCCTGTGATTAGAGATTAATTTATTAAGGTGCTCTTCTACCTGTGGTGTG 578
 DB 66145 AATTCCTGTGATTAGAGATTAATTTATTAAGGTGCTCTTCTACCTGTGGTGTG 66204

QY 579 TGTGGCGACAGCTTGAAGTGTATATAAAGAAAGAGCTCCCAATGAATCACC 637
 DB 66205 TGTGGCGAC--CAGCTAGAGGTGTATATAAAGAAAGAGCTCCCAATGAATCACC 66263

QY 638 TTATATTTATCCATTTCTATACACAGGAGCTGGAAGCAGTTTC--GAGACTTTTTCGA 695
 DB 66264 TTATATTTATCCATTTCTATACACAGGAGCTGGAAGCAGTTTCAGAGAACTTTTTCGA 66323

QY 696 TGCTTAATGTTGATCAGTTTAAAAAGATGTTTACAGTATACAAATTAAGTCAGTTTAAA 754
 DB 66324 TGCTTAATGTTGATCAGTTTAAAAAGATGTTTACAGTATACAAATTAAGTCAGTTTAAA 66382

RESULT 10
 AC067952/c 191221 bp DNA linear HTG 07-JUN-2000
 LOCUS Homo sapiens chromosome 3 clone RP11-340E6, WORKING DRAFT SEQUENCE,
 DEFINITION 34 unordered pieces.
 AC067952
 VERSION AC067952.5 GI:8569941
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 191221)
 Waterston,R.H.
 The sequence of Homo sapiens clone
 Unpublished
 JOURNAL 2 (bases 1 to 191221)
 WATERSTON,R.H.
 TITLE Direct Submission
 REFERENCE Submitted (27-APR-2000) Genome Sequencing Center, Washington
 AUTHORS University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 JOURNAL MO 63108, USA
 COMMENT On Jun 16, 2000 this sequence version replaced gi:8099097.

----- Genome Center -----
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: http://genome.wustl.edu/gsc/index.shtml
 Project Information -----
 Center project name: HNH0340E06
 Summary Statistics -----
 Sequencing vector: M13; 100%

Sequencing vector: plasmid; 0%
 Chemistry: Dye-primer ET; 100% of reads
 Chemistry: Dye-terminator Big Dye; 0% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 176290 bases at least Q40
 Consensus quality: 180539 bases at least Q30
 Consensus quality: 182588 bases at least Q20
 Insert size: 185000; agarose-fp
 Insert size: 187921; sum-of-contigs
 Quality coverage: 3.95 in Q20 bases; agarose-fp
 Quality coverage: 3.96 in Q20 bases; sum-of-contigs

----- NOTE: This is a 'working draft' sequence. It currently
 * consists of 34 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1	1361:	contig of 1361 bp in length
*	1362	gap of unknown length
*	1462	contig of 1534 bp in length
*	2996	gap of unknown length
*	3096	contig of 1939 bp in length
*	5035	gap of unknown length
*	5135	contig of 1950 bp in length
*	7085	gap of unknown length
*	7185	contig of 1518 bp in length
*	8703	gap of unknown length
*	8802	contig of 2038 bp in length
*	8803	gap of unknown length
*	10841	contig of 3091 bp in length
*	10941	gap of unknown length
*	14032	contig of 2682 bp in length
*	14131	gap of unknown length
*	14312	contig of 2029 bp in length
*	16813	gap of unknown length
*	16814	contig of 2029 bp in length
*	16914	gap of unknown length
*	18943	contig of 2778 bp in length
*	19043	gap of unknown length
*	21821	contig of 2862 bp in length
*	21921	gap of unknown length
*	24942	contig of 3022 bp in length
*	24943	gap of unknown length
*	25043	contig of 3444 bp in length
*	28487	gap of unknown length
*	28587	contig of 2862 bp in length
*	31449	gap of unknown length
*	31549	contig of 4478 bp in length
*	31549	gap of unknown length
*	36027	contig of 3229 bp in length
*	36127	gap of unknown length
*	39355	contig of 4255 bp in length
*	39356	gap of unknown length
*	39456	contig of 4255 bp in length
*	43711	gap of unknown length
*	43810	contig of 4153 bp in length
*	43811	gap of unknown length
*	47964	contig of 6235 bp in length
*	48064	gap of unknown length
*	54298	contig of 5865 bp in length
*	54399	gap of unknown length
*	60263	contig of 6254 bp in length
*	60264	gap of unknown length
*	60364	contig of 5909 bp in length
*	66178	gap of unknown length
*	66718	contig of 6080 bp in length
*	72627	gap of unknown length
*	72727	contig of 7130 bp in length
*	78807	gap of unknown length
*	78907	contig of 7541 bp in length
*	86037	gap of unknown length
*	86136	contig of 7106 bp in length
*	93678	gap of unknown length
*	93777	contig of 7106 bp in length
*	100883	gap of unknown length
*	100884	contig of 6243 bp in length
*	100984	gap of unknown length
*	107227	contig of 7251 bp in length
*	107327	gap of unknown length
*	114577	contig of 7251 bp in length

* 114578 114677: gap of unknown length
* 114678 122145: contig of 7468 bp in length
* 122146 122245: gap of unknown length
* 122246 129861: contig of 7616 bp in length
* 129862 129961: gap of unknown length
* 129962 140573: contig of 10612 bp in length
* 140574 140673: gap of unknown length
* 140674 152684: contig of 12011 bp in length
* 152685 152784: gap of unknown length
* 152785 164145: contig of 11361 bp in length
* 164146 164245: gap of unknown length
* 164246 176090: contig of 11845 bp in length
* 176091 176191: gap of unknown length
* 176191 191221: contig of 15031 bp in length.
FEATURES
source
1. 191221
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="3"
/clone="RP11-340E6"
1. 1361
/note="assembly_name:Contig22"
misc_feature
1462..2995
/note="assembly_name:Contig23"
misc_feature
3096..5034
/note="assembly_name:Contig24"
misc_feature
5135..7084
/note="assembly_name:Contig25"
misc_feature
7185..8702
/note="assembly_name:Contig26
vector_side:SP6
clone_end:SP6
vector_side:left"
8803..10840
/note="assembly_name:Contig27"
misc_feature
10941..14031
/note="assembly_name:Contig28"
misc_feature
14132..16813
/note="assembly_name:Contig29"
misc_feature
16914..18942
/note="assembly_name:Contig30"
misc_feature
19043..21820
/note="assembly_name:Contig31"
misc_feature
21921..24942
/note="assembly_name:Contig32"
misc_feature
25043..28486
/note="assembly_name:Contig33"
misc_feature
28587..31448
/note="assembly_name:Contig34"
misc_feature
31549..36026
/note="assembly_name:Contig35"
misc_feature
36127..39355
/note="assembly_name:Contig36"
misc_feature
39456..43710
/note="assembly_name:Contig37"
misc_feature
43811..47963
/note="assembly_name:Contig38"
misc_feature
48064..54298
/note="assembly_name:Contig39"
misc_feature
54399..60263
/note="assembly_name:Contig40"
misc_feature
60364..66617
/note="assembly_name:Contig41"
misc_feature
66718..72626
/note="assembly_name:Contig42"
misc_feature
72727..78806
/note="assembly_name:Contig43"
misc_feature
78907..86036
/note="assembly_name:Contig44"
misc_feature
86137..93677
/note="assembly_name:Contig45"
misc_feature
93778..100883
/note="assembly_name:Contig46"
misc_feature
100984..107226

misc_feature
/note="assembly_name:Contig47"
107327..114577
/note="assembly_name:Contig48"
misc_feature
114678..122145
/note="assembly_name:Contig49"
misc_feature
122246..129861
/note="assembly_name:Contig50"
misc_feature
129962..140573
/note="assembly_name:Contig51"
misc_feature
140674..152684
/note="assembly_name:Contig52"
misc_feature
152785..164145
/note="assembly_name:Contig53"
misc_feature
164246..176090
/note="assembly_name:Contig54"
misc_feature
176191..191221
/note="assembly_name:Contig55"
misc_feature
191221..191221
/note="assembly_name:Contig55"
BASE COUNT 52793 a 39732 c 42051 g 53340 t 3305 others
ORIGIN
Query Match 60.4%; Score 455.4; DB 2; Length 191221;
Best Local Similarity 96.5%; Pred. No. 3.1e-117;
Matches 520; Conservative 0; Mismatches 11; Indels 8; Gaps 5;
QY 223 GGGCTGGGGAATGATATATCTTCCTCCACAACTGGTGCATGTCCTGGTGGTGA 282
DB 186791 GTGCTGGGGAATGATATATCTTCCTCCACAACTGGTGCATGTCCTGGTGGTGA 186732
QY 283 CAGAACATCGCTGCCCTCTGCCAGCAGCAGCTGGTGGTCCAAAGATCGGCAATGA 342
DB 186731 CAGAACATCGCTGCCCTCTGCCAGCAGCAGCTGGTGGTCCAAAGATCGGCAATGA 186672
QY 343 GAGTGGTTGAGGCTTCTTACGCGAGTGTCCAGGCCCTGGTGGTATCTTGAATCCAG 402
DB 186671 GAGTGGTTGAGGCTTCTTACGCGAGTGTCCAGGCCCTGGTGGTATCTTGAATCCAG 186612
QY 403 TGCCTCAACAAGGCTTGAACACACACAGGAGGATGAATCTTCAATGAGAGCCGATGATC 462
DB 186611 TGCCTCAACAAGGCTTGAACACACACAGGAGGATGAATCTTCAATGAGAGCCGATGATC 186552
QY 463 TGTGCTC--TTTGACTCATCAAAAGCTTGGTTAGC--ATTGTCAGTTTATCTTCACA 518
DB 186551 TGTGCTCCTTGGGACTCATCAAAAGCTTGGTTAGCATTGTCAGTTTATCTTCACA 186492
QY 519 AATTCCTGTGATTAAGATATATTATAAGTGGTCCCTCCACTGGTGGTGG 578
DB 186491 AATTCCTGTGATTAAGATATATTATAAGTGGTCCCTCCACTGGTGGTGG 186432
QY 579 TGTGCGCACAGCTTGAAGTCTATATAAAGGAAAGAGCTCCAAATTTGATCACC- 637
DB 186431 TGTGCGGCAC-CACTTGAAGTCTATATAAAGGAAAGAGCTCCAAATTTGATCACC 186373
QY 638 TTAATAATTAACCAATTTATACACAGCAGTGAAGCAGTTTC--GAGACTTTTTCGA 695
DB 186372 TTAATAATTAACCAATTTATACACAGCAGTGAAGCAGTTTCGAGAACTTTTTCGA 186313
QY 696 TCGTTATGTTGATCATCAAGTAAAAAGATGATCAATCAATTAAGTCAAGTTTAA 754
DB 186312 TCGTTATGTTGATCATCAAGTAAAAAGATGATCAATCAATTAAGTCAAGTTTAA 186254
RESULT 11
AC060766/c 169675 bp DNA linear PRI 01-DEC-2001
LOCUS Homo sapiens chromosome 17, clone RP11-686D22, complete sequence.
DEFINITION AC060766
ACCESSION AC060766.8 GI:16931016
VERSION
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
1 (bases 1 to 169675)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 17, clone RP11-686D22
Unpublished
2 (bases 1 to 169675)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavsky,L., Bouckgalter,B., Brown,A., Burkett,G.,
Campomoro,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collymore,A., Cooke,P., DeArrelano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heatford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., Laroque,K., Lamazares,R., Landers,T., Lehocque,J.,
Levine,R., Lien,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
Meldrum,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rotman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Strange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigillo,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (20-APR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 169675)
Birren,B., Linton,L., Nusbaum,C., Lander,E., All,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Bouckgalter,B.,
Brown,A., Camarata,J., Campomoro,A., Chang,J., Chararo,B.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., DeArrelano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gird,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heatford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., Laroque,K.,
Lamazares,R., Landers,T., Lehocque,J., Levine,R., Liu,G.,
Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrum,J.,
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Roselli,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Spencer,B., Strange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigillo,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (01-DEC-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smith,A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
Project Information
Center project name: L6276
Center clone name: 686_D22
FEATURES
source
Location/Qualifiers
1..169675
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"

```

```

/map="17"
/clone="RP11-686D22"
/clone_lib="RP11 Human Male BAC"
239..259
/rpt_family="AT-rich"
263..509
/rpt_family="AluUo"
complement(618..1197)
/rpt_family="LIMC4"
1198..1218
/rpt_family="(TG)n"
1758..1933
/rpt_family="AluUo"
1934..2116
/rpt_family="LIM4"
2123..2237
/rpt_family="AluUo"
2238..2262
/rpt_family="(TAAA)n"
2263..2367
/rpt_family="AluUo"
complement(2796..3097)
/rpt_family="AluUb"
3106..3161
/rpt_family="CT-rich"
3162..3338
/rpt_family="(TTC)n"
complement(3334..3672)
/rpt_family="AluSq"
4976..5489
/rpt_family="MERS1A"
5490..5664
/rpt_family="MERS1A-int"
5656..6452
/rpt_family="MERS1A-int"
6452..6747
/rpt_family="MERS7-int"
6746..7310
/rpt_family="MERS7-int"
7314..7401
/rpt_family="LTR54B"
7402..7491
/rpt_family="MERS7-int"
9428..9535
/rpt_family="LIME"
9925..10030
/rpt_family="LIME"
10325..10367
/rpt_family="LIM4c"
10819..11180
/rpt_family="LIM4c"
11178..12391
/rpt_family="LIM4c"
complement(12392..12694)
/rpt_family="AluUo"
12695..12868
/rpt_family="LIM4c"
13176..13320
/rpt_family="LIM4c"
complement(13647..13998)
/rpt_family="LIM4c"
15861..16165
/rpt_family="AluSc"
complement(16442..16773)
/rpt_family="AluSc"
complement(17624..17785)
/rpt_family="LIM4c"
complement(17816..18567)
/rpt_family="LIME"
18588..18815
/rpt_family="AluSc"
18822..18858
/rpt_family="(CAA)n"

```

Oy	362	TAGGCGAGTTGGTCAGAGCCCTGGTGGATCTTGTAATCAGACGCCCTCAAGGCTAGAA	421
Db	93112	TAGTGCACTTGCCCTAACCCCTGGGGATCTTGTAATCAAGGTGCACAAGAGGCTGGAA	93053
Oy	422	CACACAGAGGGGATCAATTCTCTCAATAAGAGCCGATGATCTGTGCTTGGACTCATC ¹	481
Db	93052	CATTCACAGGGGATTAATTCTCTCAAATAGGACACCATGAATCTGTGCTTGGAAACTCATC	92993
Oy	482	AAGACCTTGGTTAGCAATT-CTCAGTTTTATCTTCAGAAATTCCTCTGATTAGAAGAT	540
Db	92992	AAGCAGAAATTTACAAATTTCAATCAGATTTTATTTTTCAGGAATTCCTCAAAATTAAGAAAT	92933
Oy	541	AATTTATTTAAAGGTGGTCCCTTCACCCCTGTGGTGTGTGCGGACACACAGCTTAGAG	600
Db	92932	AATTACTTAAAAACAGCTCTTCTCCTACCTCTAGAGTGTGTGCACACACACATGATTAGAG	92873
Oy	601	TGCTATATAAAAAGAGAGAGCTCCAATTTGAATCACCCTTAATTTACCAATTCCTATAC	660
Db	92872	TGCTATAGAAAAGAAAA---TTCCAAATTCGAATGACCTTCAAAATTTCCCTGTGAATAC	92816
Oy	661	AA-----CAGCAGCTGGAGACGATTTGAGACATTTTTCGATGCTTATGTTGATCAGTT	714
Db	92815	CAGGTGTGGAACAGTTCATTCATTCAGACAACTTTTGTATGCTGTGGTTGATCAGTT	92756
Oy	715	AAAAAGATGTGTACAGTAACAATTAAGTCAGTTTAAA	754
Db	92755	TTTAAAAAGGAAATGATTAATTAATTAAGTACAGTTTAAA	92716
RESULT	12		
AF142060			
LOCUS	AF142060	342 bp	mRNA linear PRI 14-JUL-1999
DEFINITION	Homo sapiens RING finger protein (ROC2) mRNA, complete cds.		
ACCESSION	AF142060		
VERSION	AF142060.1	GI:4809217	
KEYWORDS	.		
SOURCE ORGANISM	human. Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 342) Ohta,T., Michel,J.J., Schottelius,A.J. and Xiong,Y. ROC1, a homolog of APC1, represents a family of cullin partners with an associated ubiquitin ligase activity Mol. Cell 3 (4), 535-541 (1999)		
JOURNAL	MEDLINE	99247022	
PUBMED	10230407		
REFERENCE	2 (bases 1 to 342) Ohta,T., Michel,J. and Xiong,Y. Direct Submission Submitted (10-APR-1999) LIneberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Mason Farm Rd. and Manning Dr., Chapel Hill, NC 27599-7295, USA		
JOURNAL	Location/Qualifiers		
FEATURES	1..342		
source	/organism="Homo sapiens" /db_xref="taxon:9606"		
gene	1..342		
CDS	/gene="ROC2" 1..342 /note="ROC2" /note="Cullin partnering protein with associated ubiqutlin ligase activity" /codon_start=1 /product-"RING finger protein" /protein_id="AAD50147.1" /db_xref="GI:4809218" /translation="MAVDLEDGEETKCALASHSGSGSTSGDKMFSLKKVNAVAMSND VEDDTACIRVOYVMDACLROGAERKEKDQDVVWMECHNSPHNCMSLWVQNRRPCLC QQDWYVORIK"		
BASE COUNT	82 a	83 c	109 g 68 t
BRGIN			

Query Match 44.9%; Score 338.8; DB 9; Length 342;
 Best Local Similarity 99.4%; Pred. No. 6.2e-85;
 Matches 340; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGCCGACGTGGAAGACGAGAGAAACCTGGGCCCTGGCTCTCACTCCGGGAGATCA 60
 DB 1 ATGGCCGACGTGGAAGACGAGAGAAACCTGGGCCCTGGCTCTCACTCCGGGAGATCA 60

QY 61 GGCTCCAAAGTCGGGAGGCGACAAGATGTTCTCCCTCAAGAAATGGAACGGCGTGGCCATG 120
 DB 61 GGCTCCAAAGTCGGGAGGCGACAAGATGTTCTCCCTCAAGAAATGGAACGGCGTGGCCATG 120

QY 121 TGGAGCTGGGACGTGGAGTGCATAGCTGGCCATCTGCAGAGGTCCAGGTGATGGAATGCC 180
 DB 121 TGGAGCTGGGACGTGGAGTGCATAGCTGGCCATCTGCAGAGGTCCAGGTGATGGAATGCC 180

QY 121 TGGAGCTGGGACGTGGAGTGCATAGCTGGCCATCTGCAGAGGTCCAGGTGATGGAATGCC 180
 DB 121 TGGAGCTGGGACGTGGAGTGCATAGCTGGCCATCTGCAGAGGTCCAGGTGATGGAATGCC 180

QY 181 TGCTTTAGATGTCAAGCTGAAACAAACAGAGACTGTTGTGTGCTGGGAGAAATGT 240
 DB 181 TGCTTTAGATGTCAAGCTGAAACAAACAGAGACTGTTGTGTGCTGGGAGAAATGT 240

QY 241 AATCATTCTTCCACAACACTGCTGCATGTCCTGTGGGTAAACAGAAATCGCTGCCCT 300
 DB 241 AATCATTCTTCCACAACACTGCTGCATGTCCTGTGGGTAAACAGAAATCGCTGCCCT 300

QY 301 CTCTGCCAGCAGGACTGGGTGGTCCAAAGAAATGGCAAAATGA 342
 DB 301 CTCTGCCAGCAGGACTGGGTGGTCCAAAGAAATGGCAAAATGA 342

RESULT 13
 AX282592 342 bp DNA linear PAT 02-NOV-2001
 LOCUS AX282592
 DEFINITION Sequence 7 from Patent WO0175145.
 VERSION AX282592
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (sites)
 AUTHORS Issakani, S.D., Huang, J., Sheung, J. and Pray, T.R.
 TITLE Ubiquitin ligase assay
 JOURNAL Patent: WO 0175145-A7 11-OCT-2001;
 Rigel Pharmaceuticals, Inc. (us)
 FEATURES
 source 1..342
 /organism="Homo sapiens"
 /db_xref="taxon:9606"

BASE COUNT 82 a 84 c 108 g 68 t
 ORIGIN

Query Match 44.7%; Score 337.2; DB 6; Length 342;
 Best Local Similarity 99.1%; Pred. No. 1.7e-84;
 Matches 339; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGGCCGACGTGGAAGACGAGAGAAACCTGGGCCCTGGCTCTCACTCCGGGAGATCA 60
 DB 1 ATGGCCGACGTGGAAGACGAGAGAAACCTGGGCCCTGGCTCTCACTCCGGGAGATCA 60

QY 61 GGCTCCAAAGTCGGGAGGCGACAAGATGTTCTCCCTCAAGAAATGGAACGGCGTGGCCATG 120
 DB 61 GGCTCCAAAGTCGGGAGGCGACAAGATGTTCTCCCTCAAGAAATGGAACGGCGTGGCCATG 120

QY 121 TGGAGCTGGGACGTGGAGTGCATAGCTGGCCATCTGCAGAGGTCCAGGTGATGGAATGCC 180
 DB 121 TGGAGCTGGGACGTGGAGTGCATAGCTGGCCATCTGCAGAGGTCCAGGTGATGGAATGCC 180

QY 181 TGCTTTAGATGTCAAGCTGAAACAAACAGAGACTGTTGTGTGCTGGGAGAAATGT 240
 DB 181 TGCTTTAGATGTCAAGCTGAAACAAACAGAGACTGTTGTGTGCTGGGAGAAATGT 240

DB 181 TGCTTTAGATGTCAAGCTGAAACAAACAGAGACTGTTGTGTGCTGGGAGAAATGT 240

QY 241 AATCATTCTTCCACAACACTGCTGCATGTCCTGTGGGTAAACAGAAATCGCTGCCCT 300
 DB 241 AATCATTCTTCCACAACACTGCTGCATGTCCTGTGGGTAAACAGAAATCGCTGCCCT 300

QY 301 CTCTGCCAGCAGGACTGGGTGGTCCAAAGAAATGGCAAAATGA 342
 DB 301 CTCTGCCAGCAGGACTGGGTGGTCCAAAGAAATGGCAAAATGA 342

RESULT 14
 AF092877 1140 bp mRNA linear ROD 24-JUL-2001
 LOCUS AF092877
 DEFINITION Mus musculus zinc RING finger protein SAG mRNA, complete cds.
 ACCESSION AF092877
 VERSION AF092877.1 GI:4588031
 KEYWORDS
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Mus.
 REFERENCE 1 (bases 1 to 1140)
 AUTHORS Swaroop, M., Bian, J., Aviram, M., Duan, H., Bisgaier, C.L., Loo, J.A. and Sun, Y.
 TITLE Expression, purification, and biochemical characterization of SAG, a RING finger redox-sensitive protein
 JOURNAL Free Radical Biol. Med. 27, 193-202 (1999)
 REFERENCE 2 (bases 1 to 1140)
 Sun, Y.
 TITLE Alterations of SAG mRNA in human cancer cell lines: requirement for the RING finger domain for apoptosis protection
 JOURNAL Carcinogenesis. 20 (10), 1899-1903 (1999)
 REFERENCE MEDLINE 99435944
 PUBMED 10506102

QY 3 (bases 1 to 1140)
 Duan, H., Wang, Y., Aviram, M., Swaroop, M., Loo, J.A., Bian, J., Tian, Y., Mueller, T., Bisgaier, C.L. and Sun, Y.
 TITLE SAG, a novel zinc RING finger protein that protects cells from apoptosis induced by redox agents
 JOURNAL Molecular and cellular biology. 19 (4), 3145-3155 (1999)
 REFERENCE MEDLINE 99182502
 PUBMED 10082581

QY 4 (bases 1 to 1140)
 Swaroop, M., Wang, Y., Miller, P., Duan, H., Jatke, T., Madore, S.J. and Sun, Y.
 TITLE Yeast homolog of human SAG/ROC2/Rbx2/Hrt2 is essential for cell growth, but not for germination: chip profiling implicates its role in cell cycle regulation
 JOURNAL Oncogene. 19 (24), 2855-2866 (2000)
 REFERENCE MEDLINE 20309864
 PUBMED 10851089

QY 5 (bases 1 to 1140)
 Duan, H., Tsvetkov, L.M., Liu, Y., Song, Y., Swaroop, M., Wen, R., Kun, H.F., Zhang, H. and Sun, Y.
 TITLE Promotion of S-phase entry and cell growth under serum starvation by SAG/ROC2/Rbx2/Hrt2, an E3 ubiquitin ligase component: association with inhibition of p27 accumulation
 JOURNAL Molecular carcinogenesis. 30 (1), 37-46 (2001)
 REFERENCE MEDLINE 21152847
 PUBMED 11255262

QY 6 (bases 1 to 1140)
 Sun, Y.
 TITLE Direct Submission
 JOURNAL Submitted (16-SEP-1998) Department of Molecular Biology, Parke-Davis, 2800 Plymouth Rd, Ann Arbor, MI 48105, USA
 FEATURES
 source 1..1140
 /organism="Mus musculus"
 /strain="C57BL/6xCBA"
 /db_xref="taxon:10090"
 /sex="female"
 /tissue_type="lung"

```

CDS
    /dev_stage="6-8 weeks"
    17..358
    /note="redox sensitive, metal binding; expression protects
    cells from apoptosis induced by redox compounds"
    /codon_start=1
    /product="zinc RING finger protein SAG"
    /protein_id="AAD25961.1"
    /db_xref="GI:4588032"
    /translation="MADVEDGESEPCVLSHSGASGSKSGGDKMFSLKWNAAVAMSMID
    /VECDTACIRVQVWDACLRCOAENKROBDVVWMECHNSFHNCMSLMVKNRNCPLC
    QQWVVOYRIK"

BASE COUNT      302 a      224 c      287 g      327 t
ORIGIN
Query Match      44.3%; Score 334; DB 10; Length 1140;
Best Local Similarity 70.6%; Pred. No. 1.9e-83;
Matches 534; Conservative 0; Mismatches 180; Indels 42; Gaps 5;

QY  1  ATGGCCGACGTGGAAGACGAGAGAGAAACCTGGCCCTTCACCTCCGGAGCTCA 60
    |||||
Db  17  ATGGCCGACGTGGAAGACGAGAGAGAAACCTGGCCCTTCACCTCCGGAGAGCA 76
QY  61  GGCTCCAGTGGGGAGGGGGAAGATGTTCTCCCAAGAACTGGAACGGGTGGCCATG 120
    |||||
Db  77  GGCTCCAGTGGGGAGGGGGAAGATGTTCTCCCAAGAACTGGAACGGGTACCATG 136
QY  121  TGGAGCTGGAGACGTGAGTGGGATCGTGGCCATCTCAGAGGTCCAGGTGATGATGCC 180
    |||||
Db  137  TGGAGCTGGAGACGTGAGTGGGATCGTGGCCATCTCAGAGGTCCAGGTGATGATGCC 196
QY  181  TGTCTTAGATGTCAAGCTGAAAAACAAGAGACTGTGTTGTGTGGGGAAGATGT 240
    |||||
Db  197  TGCCTTCGATGTCAAGCTGAAAAACAAGAGAGACTGTGTTGTGTGGGGAAGATGT 256
QY  241  AATCATTCCTCCCACTGCTGCATGTCCCTGTGTGGTGAACAACAATGCTGCGCT 300
    |||||
Db  257  AACAATTCCTCCCACTGCTGCATGTCCCTGTGTGGTGAACAACAATGCTGCGCT 316
QY  301  CTCTGCCAGCAGAGACTGGGTGGTCCAAAGAAATCGGCAATGAGAGTGGTTAGAGGCTTC 360
    |||||
Db  317  CTCTGCCAGCAGAGACTGGGTGGTCCAAAGAAATCGGCAATGAGAGTGGTCCAGCGCTC 376
QY  361  TTAGCGCAGTTGTTCAGAGCCCTGTGTGATCTTGTAAATCCAGTGCCCTACAAAGGCTAGA 420
    |||||
Db  377  CTGGTGTGTGCT-----GACCCTGGACAAAGACTAA 409
QY  421  ACACCTACGSGGATGATCTTCAATATAGAGCCATGATCTGTGTCTTTGGACTAT 480
    |||||
Db  410  ACACCTACGSGGATGATCTTCAATATAGAGAG--AGAGATGATCTGTGGCTTGGAGACTAC 467
QY  481  CAAAGCCTGGT--AGCATTTGTCAGTTTATCTTCAAAATTCCTGTGTAAAGAG 538
    |||||
Db  468  CAAAGCCTGGTATTTATTTATTTGTCTGTGTATTGGGAATTTCTCAAAATTT--AAG 524
QY  539  ATAATTTTAAAGGTGCTCTTCTTACCTCTGTGTGTGTGTGTGCGGACACAGCTTGA 598
    |||||
Db  525  ATAATTTTAAAGGTGCTCTTCTTACCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 576
QY  599  AGTGTATTAATAAAGGAAGAGCTCCAATTGATCACCCTTATTAATTAACCATTTCTAT 658
    |||||
Db  577  ATGCATTAAGAGGCGAGAACACACAGAAATATGATCTTTGTATTCGTATACCCACGACTGA 636
QY  659  ACAACAGCAGTGGAGAGAGTTTCAGACTTTTGGATGCTTATAGGTGATGATGATTAATA 718
    |||||
Db  637  ACATTTGTCTTACACAGAACATTTTGTGTGTATGTCTGAGGGTTAAAAAATATGATTA 696
QY  719  AAGATGTTACAGTAAACAATAAAGTCAAGTTTAA 754
    |||||
Db  697  ACGATGTTACAGTAAACAATAATGATTAATAA 732

```

LOCUS	BC011127	1125 bp	mRNA	linear	ROD 30-JUL-2001
DEFINITION	Mus musculus, similar to ring finger protein 7, clone MGC:19076				
ACCESSION	IMAGE:4194107, mRNA, complete cds.				
VERSION	BC011127				
KEYWORDS	BC011127.1 GI:15029807				
SOURCE	MGC.				
ORGANISM	house mouse. Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 1125)				
AUTHORS	Strausberg, R.				
TITLE	Direct Submission				
JOURNAL	Submitted (25-JUL-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA				
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov				
COMMENT	Contact: MGC help desk Email: cgaps-remail.nih.gov Tissue Procurement: Jeffrey E. Green, M.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center Center code: BCM-HGSC Web site: http://www.hgsc.bcm.tmc.edu/cdna/ Contact: villalobom@bcm.tmc.edu . Villalobon, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia, A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W., Muzny, D.M., Gibbs, R.A.				
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAC Plate: 24 Row: K Column: 5 This clone was selected for full length sequencing because it passed the following selection criteria: genomeScan gene prediction, similarity but not identity to protein.				
SOURCE	Location/Qualifiers				
	1..1125				
	/organism="Mus musculus"				
	/db_xref="taxon:10090"				
	/clone="MGC:19076 IMAGE:4194107"				
	/tissue_type="Liver, normal. 5 month old male mouse."				
	/clone_lib="NCI CGAP_L19"				
	/lab_host="DH10B"				
	/note="Vector: pCMV-SPORT6"				
	12..353				
	/codon_start=1				
	/product="similar to ring finger protein 7"				
	/protein_id="AAH11127.1"				
	/db_xref="GI:15029808"				
	/translation="MADVDEDEEPCVLSHSIGSGKSGDKRSLKKNVAVAMRSDVCEQTCICRQVMDACLRCQAEKRKQEDCVYVWGECHNSHFNCMSLWVONRNCPLCQODWVVOIRIG"				
	QODWVVOIRIG"				
BASE COUNT	304 a	219 c	286 g	316 t	
ORIGIN					
Query Match	44.2%	Score 333.4	DB 10	Length 1125	
Best Local Similarity	71.3%	Pred. No. 2.8e-83			
Matches 540	Conservative 0	Mismatches 176	Indels 41	Gaps 6	
OY	1	ATGCCGACGTRGAGAGCGAGAGAAACCTGCGCCCTGCGCTCATCTCCGGAGCTCA	60		
Db	12	ATGCCGACGTRGAGAGCGAGAGAAACCTGCGCTCATCTCCGAGCGCA	71		
OY	61	GACCTCAAGTCGGGAGCGGACAAAGTGTTCGCCCAAGAGTGAACGGGTGCCA	120		
Db	72	GACCTCAAGTCGGGAGCGGACAAAGTGTTCCTTCACAAAGAGTGAACGGGTGCCA	131		
OY	121	TGAGACTTGAGACGTRGAGATGCATACGTCGCGCATCTCGAGGCTCAGGTGATGATGCC	180		

```

Db      132  TGGAGCTGGGACGTTGAGTGAGTACCTGTCATCTGCAGGGTCCAGTGATGATGCC 191
QY      181  TGTCTTAGATGTCAAGTGTGAACCAACAGAGAGACTGTGTGTGTGTGTGGGAGAAATGT 240
Db      192  TGCCTTGATGTCAAGTGTGAACCAACAGAGAGACTGTGTGTGTGTGTGGGAGAGTGT 251
QY      241  AATCATTCCTTCCACAACTGCTGCATGTCCCTGTGGGTGAACAGAACAAATCGTGCCT 300
Db      252  AACCATTCCTTCCACAACTGCTGCATGTCCCTGTGGGTGAACAGAACAAATCGTGCCT 311
QY      301  CTCGCCGAGAGAGCTGGGTGTCCAAAGAAATGAGAGATGAGTGTAGAGGCTTC 360
Db      312  CTGTGCCAGAGAGCTGGGTGTCCAAAGAAATGAGAGATGAGTGTAGAGGCTTC 371
QY      361  TTAGCGCAGTTGTTCAGAGCCCTGTGTGATCTGTAAATCCAGTGCCTTACAAAGGCTAGA 420
Db      372  CTGGTGTGTTG-----CTGACCCCTGGACAAAGACTAA 404
QY      421  ACACATACAGGGAGATCTTCAAAATAGAGCCGATGATCTGTGTGTGTGTGTGTGTGT 480
Db      405  ACACATGACAGGGAGATCTTCAAAATAGAGCCGATGATCTGTGTGTGTGTGTGTGTGT 462
QY      481  CAAAGCCTTGGT--AGCATTTGTCAGTTTATCTTCAAAATCTCTGTGATTAGAGAG 538
Db      463  CAAAGCCTTGGT--AGCATTTGTCAGTTTATCTTCAAAATCTCTGTGATTAGAGAG 519
QY      539  ATAATTTAATTAAGGTGCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 597
Db      520  ATAATTTAATTAAGGTGCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 579
QY      598  AAGTCTTAAATAAGGAAGGCTCCAAATGATCACTTAATTAATTAATTAATTAATTAATTA 657
Db      580  AAGACCGAGAACACCAAAATGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 633
QY      658  TACAACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 717
Db      634  AACATTTGTTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 693
QY      718  AAAGAATGTTACAGTACAAATAAAGTGCAGTTTAA 754
Db      694  AACGAATGTTACAGTACAAATAAAGTGCAGTTTAA 730

```

Search completed: July 25, 2002, 19:14:30
 Job time: 18349 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 25, 2002, 17:22:57 ; Search time 312.78 Seconds
(without alignments)
6257.697 Million cell updates/sec

Title: US-09-509-779-1

Perfect score: 1140
Sequence: 1 GTTCTGGCGCGCCCATG.....ACTAATTCATCAATTAATG 1140

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3412872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : N.Geneseq_032802.*
1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1140	100.0	1140	20	Mouse sensitive to
2	364.4	32.0	962	21	Human ORFX ORF3059
3	363.4	31.9	836	22	Human CDNA encodin
4	362.6	31.8	836	22	Human CDNA encodin
5	334	29.3	754	20	Human sensitive to
6	332.4	29.2	754	20	Human sensitive to
7	332.4	29.2	754	20	Human sensitive to
8	332.4	29.2	754	20	Human sensitive to
9	332.4	29.2	754	20	Human sensitive to

10	332.4	29.2	754	20	AAH87324	Human sensitive to
11	332.4	29.2	754	20	AAH87325	Human sensitive to
12	332.4	29.2	754	20	AAH87327	Human sensitive to
13	332.4	29.2	754	20	AAH87328	Human sensitive to
14	332.4	29.2	754	20	AAH87331	Human sensitive to
15	330.8	29.0	754	20	AAH87332	Human sensitive to
16	330.8	29.0	754	20	AAH87332	Human sensitive to
17	330.8	29.0	754	20	AAH87332	Human sensitive to
18	330.8	29.0	754	20	AAH87329	Human sensitive to
19	330.8	29.0	754	20	AAH87330	Human sensitive to
20	327.6	28.7	754	20	AAH87326	Human sensitive to
21	310	27.2	747	20	AAH87313	Human sensitive to
22	303.6	26.6	342	21	AAH8883	Nucleotide sequenc
23	303.6	26.6	342	22	AAH43570	ROC2 coding sequen
24	232.8	20.4	706	20	AAH87316	Human sensitive to
25	223.6	19.6	441	21	AAH77493	Human ORFX ORF3048
26	200.4	16.6	1152	22	AAH25847	Human apoptosis as
27	182.6	16.0	224	20	AAH41085	Human secreted pro
28	111.6	9.8	596	22	AAH25052	Human breast cance
29	108.2	9.5	169	20	AAH41482	Human secreted pro
30	102.6	9.0	630	22	AAH22339	Human breast cance
31	90.6	7.9	327	21	AAH96882	Nucleotide sequenc
32	90.6	7.9	482	21	AAH03896	Human secreted pro
33	90.6	7.9	508	21	AAH74978	DNA encoding a hum
34	90.6	7.9	5347	24	AAH94844	Human DNA sequence
35	87.8	7.7	498	23	ABL14673	Drosophila melanog
36	87.8	7.7	10025	23	ABL14672	Drosophila melanog
37	84.8	7.4	557	22	AAH97861	Marine 7-transmemb
38	84	7.4	560	23	ABL18585	Drosophila melanog
39	84	7.4	2560	23	ABL18584	Drosophila melanog
40	84	7.4	3038	23	ABL02194	Drosophila melanog
41	84	7.4	4760	23	ABL20826	Drosophila melanog
42	84	7.4	4764	23	ABL11922	Drosophila melanog
43	83.4	7.3	3208	22	AAH12859	Human novel protei
44	81.8	7.2	692	23	ABL22527	Drosophila melanog
45	79.4	7.0	504	21	AAH74980	DNA encoding a mur

ALIGNMENTS

RESULT 1	
AAH87313	
ID	AAH87313 standard; CDNA: 1140 BP.
AC	AAH87313;
DT	27-SEP-1999 (first entry)
DE	Mouse sensitive to apoptosis (SAG) gene.
KW	SAG gene; sensitive to apoptosis; mouse; cancer; tumour;
KW	neurodegenerative disease; muscular dystrophy; wound healing;
KW	vulnerary; therapy; ds.
OS	Mus musculus.
FT	Key
FT	CDS
FT	Location/Qualifiers
FT	17..358
FT	/*tag= a
PN	MO9932514-A2.
PD	01-JUL-1999.
PF	15-DEC-1998; 98WO-US26705.
PR	11-SEP-1998; 98US-0099840.
PR	19-DEC-1997; 97US-0068179.
PA	(WARN) WARNER LAMBERT CO.
PI	Sun Y;

PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226868.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232387.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236337.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 13-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.

PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 (HUMA-) HUMAN GENOME SCI INC.
 Rosen CA, Barash SC, Ruben SM;
 WPI: 2001-488783/53.
 P-PSDB; AAU16332.
 New nucleic acid molecules encoding 461 human secreted proteins for
 diagnosing, preventing, treating or ameliorating medical conditions and
 used as food additives or preservatives -
 PS Claim 1; SEQ ID No 498; 980pp; English.
 The invention relates to isolated nucleic acid molecules and their
 encoded secreted proteins. The nucleic acids and proteins are used to
 prevent, treat or ameliorate a medical condition in e.g. humans, mice,
 rabbits, goats, horses, cats, dogs, chickens or sheep. They
 are also used in diagnosing a pathological condition or susceptibility
 to a pathological condition. Antibodies to the proteins can also
 be used in alleviating symptoms associated with the disorders and in
 diagnostic immunoassays e.g. radioimmunoassays or enzyme linked
 immunosorbant assays (ELISA). Disorders which are diagnosed or treated
 include autoimmune diseases e.g. rheumatoid arthritis,
 hyperproliferative disorders e.g. neoplasms of the breast or liver,
 cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
 e.g. cerebral ischemia, angiogenesis, nervous system disorders e.g.
 Alzheimer's disease, infections caused by bacteria, viruses and fungi
 and ocular disorders e.g. corneal infection, and many other
 disorders listed in the specification. The polypeptides can also
 be used to aid wound healing and epithelial cell proliferation, to
 prevent skin aging due to sunburn, to maintain organs before

Query Match	31.9%;	Score 363.4;	DB 22;	Length 836;
Best Local Similarity	78.1%;	Pred. No. 6.2e-94;		
Matches 502; Conservative	0;	Mismatches 106;	Indels 35;	Gaps 4

KW	corneal infection; wound healing; epithelial cell proliferation,
KM	skin ageing; food additive; preservative; antiproliferative.
XX	
OS	Homo sapiens.
XX	
PN	W0200155322-A2.
PD	
XX	
XX	
PF	17-JAN-2001; 2001WO-US01341.
XX	
PR	31-JAN-2000; 2000US-0179065.
PR	04-FEB-2000; 2000US-0180628.
PR	24-FEB-2000; 2000US-0184664.
PR	02-MAR-2000; 2000US-0186350.
PR	16-MAR-2000; 2000US-0189874.
PR	17-MAR-2000; 2000US-0190076.
PR	18-APR-2000; 2000US-0198123.
PR	19-MAY-2000; 2000US-0205315.
PR	07-JUN-2000; 2000US-0209467.
PR	28-JUN-2000; 2000US-0214886.
PR	30-JUN-2000; 2000US-0215135.
PR	07-JUL-2000; 2000US-0216647.
PR	07-JUL-2000; 2000US-0216880.
PR	11-JUL-2000; 2000US-0217487.
PR	11-JUL-2000; 2000US-0217496.
PR	14-JUL-2000; 2000US-0218290.
PR	26-JUL-2000; 2000US-0220963.
PR	26-JUL-2000; 2000US-0220964.
PR	14-AUG-2000; 2000US-0224518.
PR	14-AUG-2000; 2000US-0224519.
PR	14-AUG-2000; 2000US-0225213.
PR	14-AUG-2000; 2000US-0225214.
PR	14-AUG-2000; 2000US-0225266.
PR	14-AUG-2000; 2000US-0225267.
PR	14-AUG-2000; 2000US-0225268.
PR	14-AUG-2000; 2000US-0225270.
PR	14-AUG-2000; 2000US-0225447.
PR	14-AUG-2000; 2000US-0225757.
PR	14-AUG-2000; 2000US-0225758.
PR	14-AUG-2000; 2000US-0225759.
PR	18-AUG-2000; 2000US-0226279.
PR	22-AUG-2000; 2000US-0226681.
PR	22-AUG-2000; 2000US-0226688.
PR	22-AUG-2000; 2000US-0227182.
PR	23-AUG-2000; 2000US-0227009.
PR	30-AUG-2000; 2000US-0228924.
PR	01-SEP-2000; 2000US-0229287.
PR	01-SEP-2000; 2000US-0229287.
PR	01-SEP-2000; 2000US-0229343.
PR	01-SEP-2000; 2000US-0229344.
PR	01-SEP-2000; 2000US-0229345.
PR	05-SEP-2000; 2000US-0229509.
PR	05-SEP-2000; 2000US-0229513.
PR	06-SEP-2000; 2000US-0230437.
PR	06-SEP-2000; 2000US-0230438.
PR	08-SEP-2000; 2000US-0231242.
PR	08-SEP-2000; 2000US-0231243.
PR	08-SEP-2000; 2000US-0231244.
PR	08-SEP-2000; 2000US-0231313.
PR	08-SEP-2000; 2000US-0231414.
PR	08-SEP-2000; 2000US-0232080.
PR	08-SEP-2000; 2000US-0232081.
PR	12-SEP-2000; 2000US-0231668.
PR	14-SEP-2000; 2000US-0232397.
PR	14-SEP-2000; 2000US-0232398.
PR	14-SEP-2000; 2000US-0232399.
PR	14-SEP-2000; 2000US-0232400.
PR	14-SEP-2000; 2000US-0232401.
PR	14-SEP-2000; 2000US-0233063.
PR	14-SEP-2000; 2000US-0233064.
PR	14-SEP-2000; 2000US-0233065.
PR	21-SEP-2000; 2000US-0234223.
PR	21-SEP-2000; 2000US-0234223.
PR	21-SEP-2000; 2000US-0234274.

DB 425 caaaggctagaacacacacacagggagatgaatcttcaaatggagccgatctggtc 484
 QY 455 CTTTGAGACTCACCAAGGCTTCTTATTATTGTTGTTGGAAATTCCTC 514
 DB 485 cttgggaactcaacaagccttggttagcatlcttgcaatcttcagaaattc 544
 QY 515 TACAATT---AAGTAAATTGTTAAAGGCTTCCCTACCTGCTGTGTGTGA 571
 DB 545 tggatlaagaagaataattatcaagaagtgcttccttcctctgtgtgtcgcg 604
 QY 572 TACGAATGCATAGAGACGAGACACAGAAATGATC 610
 DB 605 cac-acagcttagaagtgctataaaaaaagaaagagctc 642

RESULT 5
 AAX87314
 ID AAX87314 standard; cDNA: 754 BP.
 AC AAX87314;
 DT 27-SEP-1999 (first entry)
 DE Human sensitive to apoptosis (SAG) gene.
 XX
 KW SAG gene; sensitive to apoptosis; human; cancer; tumour;
 KW neurodegenerative disease; muscular dystrophy; wound healing;
 KW vulnary; therapy; ds.
 XX
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 1..342
 FT /tag a
 PN W0932514-A2.
 PD 01-JUL-1999.
 PE 15-DEC-1998; 98MO-US26705-
 PR 11-SEP-1998; 98US-0099840.
 PR 19-DEC-1997; 97US-0068170-
 XX
 PA (WARN) WARNER LAMBERT CO.
 PI Sun Y;
 DR WPI: 1999-430152/36.
 DR P-PSDB; AAY06492.
 PT SAG: Sensitive to Apoptosis Gene and related proteins, useful for
 PT promoting cell growth and protecting cells against apoptosis
 XX
 PS Claim 9; Page 50-51; 84pp; English.
 XX
 CC This is the nucleotide sequence of human sensitive to apoptosis
 CC (SAG) cDNA, which codes for a novel redox-sensitive, haem-binding
 CC protein (see AAY06492) with a zinc RING finger domain that promotes
 CC cell growth, protects cells from apoptosis, scavenges oxygen
 CC radicals and can be used for the reversion of a tumour phenotype.
 CC The cDNA was isolated from a HeLa cell cDNA library using mouse
 CC SAG cDNA (see AAX87313) as probe. SAG is highly conserved among
 CC species. Disruption in yeast was shown to be lethal. SAG deletion
 CC mutants (see AAX87315-16) have been identified in human cancer lines,
 CC suggesting a role in carcinogenesis. SAG genes, and mutant SAG
 CC genes, can be used to protect cells from apoptosis induced by redox
 CC reagents. Antisense SAG genes can be used to inhibit the growth of
 CC tumour cells. The SAG genes can also be used for the recombinant
 CC production of the SAG proteins. The SAG proteins can be used to
 CC scavenge oxygen radicals in organisms and to promote wound healing.
 CC Additionally, the SAG genes or their complements can be used to

CC promote or inhibit the growth of plant cells (all claimed). The SAG
 CC protein is also an ideal molecular target in the development of
 CC drugs against neurodegenerative disorders, cancers and muscle
 CC dystrophy.
 XX
 SQ Sequence 754 BP; 205 A; 155 C; 201 G; 193 T; 0 other;

Query Match 29.3%; Score 334; DB 20; Length 754;
 Best Local Similarity 70.6%; Pred. No. 1,6e-85;
 Matches 534; Conservative 0; Mismatches 180; Indels 42; Gaps 5;

QY 17 ATGGCCGAGCTGAGAGACGGCGAGAAACCTGCTCTTCTCGGACACCGGAGCGCA 75
 DB 1 atggccgagctggaagacggaaggaacccgcgccttgccctcaaccgcggaagctca 60
 QY 77 GGCTCCAGTCGGAGAGCGACGAAGATGTTCTCTCAGAGATGGAACGGGTAGCCATG 136
 DB 61 ggtccaaagtcggagagcgacgaagatgtctccctcaagaagatggaagcgggtgcac 120
 QY 137 TGGAGCTGGAGCTTGAGTCCGATACCTGTCATCTGCAGGGTCCAGGTGATGATGCC 196
 DB 121 tggagctggagacgltgagtgatgacgtgcacatctgcaggggtccaggtgataagc 180
 QY 197 TGCCTTCGATGTCAGAGCTGGAACAAGCAAGAGAGAGAGCTGTGTGTGCTGGGAGAGT 236
 DB 181 tgccttagatgltcaagctgaaacaaagagactgltgtgtgtcgtgggaagatgt 240
 QY 257 AACCAATTCCTCCACAACCTGCTCATGTCCTGTGGTGAACAAGAACATTCCTGCCCT 316
 DB 241 aatcaatctccacacactctgcatgtccctgtgtgtgtgaagacagacacatcgtccct 300
 QY 317 CTGTGCAGAGAGACAGTCGAGTCAAGTCAAGTGGCAAGTGAAGTGGCCAGGCGCTC 376
 DB 301 ctgtgcagagagactgt 360
 QY 377 CTGCTGTGCTT-----CTGACCTGTGCAACAAGACTAA 409
 DB 361 ttgagcagltgttcaagagccctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 420
 QY 410 ACACTGCAAGGGATTCATCCTTGAGAGAG--AAGAGATGCTGTGGCCTTGGACTCAC 467
 DB 421 acaactcaagggatgaattctcaaatagaagccgagatcgtgtgtgtgtgtgtgtgt 480
 QY 468 CAAGAGCTTCTTATTAATTGCTGTAGTTTGGGAATTCCTACAATT---AAG 524
 DB 481 caaagccttggttag--cattgtcagtttattcttcaagaattctctgtgattaaga 538
 QY 525 ATAAATTTGTTAAAGATGCTTCTACCTGCTGTGTGTGT-----GTGATACGA 576
 DB 539 ataatttataaaggt 598
 QY 577 ATGCATAGAGAGACGAGACACAGAAATGATCTTTGTTATCTGTAACCCAGACTGA 636
 DB 599 agtgcataaaaaaagaagaagctccaattgataaccattatattaccatttcat 658
 QY 637 ACATTTGTTACAGAGAAGACATTTGTTGTTTATCTTACGAGGTTAAATAAGTAA 696
 DB 659 acaacagggcgtggaagcagtttcgagacttllcgatgtctatgtgtgacttaaa 718
 QY 697 ACGAATGTACAGTACCAATAAATGCAATTGAA 732
 DB 719 aagaatgttcaagtaacaataaagaagtcagtttaa 754

RESULT 6
 AAX87317
 ID AAX87317 standard; cDNA: 754 BP.
 AC AAX87317;
 DT 27-SEP-1999 (first entry)
 XX

19-DBC-1997; 97US-0066179.
(WARN) WARNER LAMBERT CO.
Sun Y;
WPI; 1999-430152/36.
P-PSDB; AAY06496.
SAG: Sensitive to Apoptosis Gene and related proteins, useful for promoting cell growth and protecting cells against apoptosis
Claim 15; Page 60-61; 84pp; English.
This is the nucleotide sequence of human sensitive to apoptosis (SAG) mutant gene MM2, which codes for a SAG protein (see AAY06496) in which the Cys residue at position 53 of the native protein (see AAY06492) is replaced by a Ser residue owing to a mutation of codon 50 from TGC to AGC obtained by site-directed mutagenesis of SAG cDNA. This residue is at a haem binding site of SAG. Single and double SAG mutants (see AAX87317-31) were made in order to determine the role of each cysteine residue of SAG in haem binding and SAG oligomerization. These properties were unaffected by the MM2 mutation. SAG is a novel zinc finger protein that promotes cell growth, protects cells from apoptosis, scavenges oxygen radicals and can be used for the reversion of a tumour phenotype. SAG genes, and mutant SAG genes, can be used to protect cells from apoptosis induced by redox reagents. They can also be used for the recombinant production of SAG proteins, which are molecular targets in the development of drugs against neurodegenerative disorders, cancers and muscle dystrophy, and promoting wound healing.

Db	481	caaaagccttgtag--cattgacgtttattcttcagaaatctctctgattaagaag	538
Oy	525	ATTAATTGTTAAAAATGGCCCTTTTCCTAACCTCTGCTGTGTGTGT-----GTGATACGA	576
Db	539	ataatttatttlaaagtgtgcttccttactctctgtggtgtgtgtgcgcacacagcttaga	598
Oy	577	ATGCATTAGACAGCGAGAACACAGAAATAATCATTTTGTATTATCTGTACCCAGACATGGA	636
Db	559	agtcgtataaaaaaaggaagaagcctcccaattgtaatcactataattaccatttctat	658
Oy	637	ACATTGTGTTCCACGAGAACACATTGTTGTGTTTATGCTTGAGCGTTAAAAAATAGATAA	696
Db	659	acaaacagcgagtggaagcagcttcgagacttcttcgactgtctatgcttgatcagttaaa	718
Oy	697	ACGAATGTTTACGTACACAAATTAATGCATTGAAAA 732	
Db	719	aagaatgtctacagtaacaataaaagtgcagtttaa 754	
RESULT 8			
AAx87320			
ID	AAx87320 standard; cDNA; 754 BP.		
xx			
xx	AAx87320;		
xx			
xx	27-SEP-1999 (first entry)		
xx			
xx	Human sensitive to apoptosis (SAG) gene mutant MM4.		
xx			
xx	SAG gene: sensitive to apoptosis; human; cancer; tumour;		
KW	neurodegenerative disease; muscular dystrophy; wound healing;		
KW	vulnerable; therapy; mutant; ds.		
xx			

Query Match	29.28;	Score 332.4;	DB 20;	Length 754;
Best Local Similarity	70.58;	Pred. No. 4.7e-85;		
Matches 533; Conservative	0;	Mismatches 181;	Indels 42;	Gaps 5

[illegible]

FT	mutation	
FT		/replace(181,T)
FT		/*tag= b
FT		/note= "C61S mutation"
XX		
PN	M09932514-A2.	
XX		
PD	01-JUL-1999.	
XX		
PF	15-DEC-1998;	98WO-US26705.
XX		
PR	11-SEP-1998;	98US-0099840.
PR	19-DEC-1997;	97US-0068179.
XX		
PA	(WARNER) WARNER LAMBERT CO.	
XX		
P1	Sun Y;	
XX		
DR	WPI: 1999-430152/36.	
DR	P-PSDB; AAY06498.	
XX		
PT	SAG: Sensitive to Apoptosis Gene and related proteins, useful for	
PT	promoting cell growth and protecting cells against apoptosis	
XX		
PS	Claim 15, Page 64-65; 84pp; English.	
XX		
CC	This is the nucleotide sequence of human sensitive to apoptosis	
CC	(SAG) mutant gene MM4, which codes for a SAG protein (see AAY06498)	
CC	in which the Cys residue at position 61 of the native protein (see	
CC	AAY06492) is replaced by a Ser residue owing to a mutation of codon	
CC	50 from TGC to AGC obtained by site-directed mutagenesis of SAG	
CC	cDNA. This residue is at zinc-finger finger 1 of SAG. Single and	
CC	double SAG mutants (see AAX87317-31) were made in order to determine	
CC	the role of each cysteine residue of SAG in haem binding and SAG	
CC	oligomerization. These properties were unaffected by the MM4	

CC mutation. SAG is a novel zinc finger protein that promotes
 CC cell growth, protects cells from apoptosis, scavenges oxygen
 CC radicals and can be used for the reversion of a tumour phenotype.
 CC SAG genes, and mutant SAG genes, can be used to protect cells from
 CC apoptosis induced by redox reagents. They can also be used for the
 CC recombinant production of SAG proteins, which are molecular targets
 CC in the development of drugs against neurodegenerative disorders,
 CC cancers and muscle dystrophy, and promoting wound healing.
 XX
 XX
 SQ Sequence 754 BP; 206 A; 155 C; 201 G; 192 T; 0 other;

Query Match 29.2%; Score 332.4; DB 20; Length 754;
 Best Local Similarity 70.5%; Pred. No. 4.7e-85;
 Matches 533; Conservative 0; Mismatches 181; Indels 42; Gaps 5;

OY 17 ATGGCCGACGTGGAGAGCGGACGACACCCCTGCTCTTCTTCCGACTCCGGAGCGCA 76
 DB 1 atggccgacgtggaagagcgagaaacctgagccctgcctcaccacgcggagctca.60
 OY 77 GGCTCCAGTCCGGAGCGGACGACAGATGTTCTCTCAAGAACTGGAGCGGTTAGCCATG 136
 DB 61 ggctccagtcggagcgagcgacaaagatgtctccccaagaatggaacgcytgccatg 120
 OY 137 TGGAGCTGGAGACGTTGAGTGCATCTGTGCATCTGCAGGTCGAGGTGATGATGATGCC 196
 DB 121 tggagctggagacgtgagtgagatagtcgcatctcgaggttcaagtgatgtagtgcc 180
 OY 197 TGCCTTCGATGTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGT 256
 DB 181 tgccttcgatgtcaagcaagcaagcaagcaagcaagcaagcaagcaagcaagcaagt 240
 OY 257 AACCATTCCTTCACAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 316
 DB 241 aaccattccttcacaaactgctgctgctgctgctgctgctgctgctgctgctgctgct 300
 OY 317 CTGTCCGACGAGACTGGGTAGTCCAAAGATGCGCAATGAGAGGTGCCAGCGCCTC 376
 DB 301 ctgtccgacgagactgggtagtcctcaagaatcgcaaatgagatgtagaagcttc.360
 OY 377 CTGTTGTGGTTG-----CTGACCCCTGGACAAACACTTAA 409
 DB 361 ttgagcgacgttgttcagagccctgtgtgactctgttaatccagtcgccacaagaagctaga 420
 OY 410 ACATGCGACGGGATTCATCTTGAGAGAG--AGAGATGCTGTGCGCCTTTGAGACTGAC 467
 DB 421 acatcgacgggatcatcttcaatagagcgatgactgtgtcttggactcat 480
 OY 468 CAAAGGCTTGCTTTTATTTGTCTGTTTGTGGAATTTCTTCACAAT---AAG 524
 DB 481 caaagccttggttag--cattgtcagtttactctcagaatctctgtgtttagaagag 538
 OY 525 ATAAATTTGTTAAATAGCCTTCTTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 576
 DB 539 ataattttaaagtgctctcctcactcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 598
 OY 577 ATGATGAAGAAGCGAGACACGACGAAATGATCTTTGTTTATCTGTAACCCAGCATGGA 636
 DB 599 atgataaagaagaagaagatccaatgaatccaattataattaccattctat 658
 OY 637 ACATGTGTTTCCACAGAGACATTTGTGTTTATGCTTGAGGTTTAAATAATGATTA 696
 DB 659 acaacagcgagtggaagcagtttcgaagacttttcgatgtctatgtatcagttaaaa 718
 OY 697 ACGAATGTTTACAGTAACAATTAATGATTGAAAA 732
 DB 719 aagaatgttaccagtaacaataaagtcagtttaa 754

RESULT 9
 ID AAX87321
 XX AAX87321 standard; cDNA; 754 BP.

AC AAX87321;
 XX 27-SEP-1999 (first entry)
 XX
 XX Human sensitive to apoptosis (SAG) gene mutant MM5.
 DE
 XX SAG gene; sensitive to apoptosis; human; cancer; tumour;
 KW neurodegenerative disease; muscular dystrophy; wound healing;
 KW vulnerability; therapy; mutant; ds.
 XX
 XX Homo sapiens.
 OS Synthetic.
 XX
 XX Key location/Qualifiers
 FH 1..342
 FT CDS /*tag= a
 FT mutation /*tag= b
 FT /*note= "C80S mutation"
 XX
 XX W09932514-A2.
 XX
 XX 01-JUL-1999.
 XX
 XX 15-DEC-1998; 98WO-US26705.
 XX
 XX 11-SEP-1998; 98US-0099840.
 XX 19-DEC-1997; 97US-0068179.
 XX (WARN) WARNER LAMBERT CO.
 XX
 XX Sun Y;
 PI WPI: 1999-430152/36.
 DR P-PSDB: AAY06499.
 XX
 XX SAG: Sensitive to Apoptosis Gene and related proteins, useful for
 PT promoting cell growth and protecting cells against apoptosis
 PT
 PS Claim 15; Page 66; 84pp; English.
 XX
 CC This is the nucleotide sequence of human sensitive to apoptosis
 CC (SAG) mutant gene MM5, which codes for a SAG protein (see AAY06499)
 CC in which the Cys residue at position 80 of the native protein (see
 CC AAY06492) is replaced by a Ser residue owing to a mutation of codon
 CC 80 from TGC to AGC obtained by site-directed mutagenesis of SAG
 CC cDNA. This residue is in zinc-ring finger 1 of SAG. Single and
 CC double SAG mutants (see AAX87317-31) were made in order to determine
 CC the role of each cysteine residue of SAG in haem binding and SAG
 CC oligomerization. These properties were unaffected by the MM5
 CC mutation. SAG is a novel zinc finger protein that promotes
 CC cell growth, protects cells from apoptosis, scavenges oxygen
 CC radicals and can be used for the reversion of a tumour phenotype.
 CC SAG genes, and mutant SAG genes, can be used to protect cells from
 CC apoptosis induced by redox reagents. They can also be used for the
 CC recombinant production of SAG proteins, which are molecular targets
 CC in the development of drugs against neurodegenerative disorders,
 CC cancers and muscle dystrophy, and promoting wound healing.
 CC
 XX
 SQ Sequence 754 BP; 206 A; 155 C; 201 G; 192 T; 0 other;

Query Match 29.2%; Score 332.4; DB 20; Length 754;
 Best Local Similarity 70.5%; Pred. No. 4.7e-85;
 Matches 533; Conservative 0; Mismatches 181; Indels 42; Gaps 5;

OY 17 ATGGCCGACGTGGAGAGCGGACGACACCCCTGCTCTTCTTCCGACTCCGGAGCGCA 76
 DB 1 atggccgacgtggaagagcgagaaacctgagccctgcctcaccacgcggagctca 60
 OY 77 GGCTCCAGTCCGGAGCGGACGACAGATGTTCTCTCAAGAACTGGAGCGGTTAGCCATG 136
 DB 61 ggctccagtcggagcgagcgacaaagatgtctccccaagaatggaacgcytgccatg 120

OY		137	TGGAGCTGGGACGTTTGGATGGGAATACCCTGTGCCATCTCGACAGGGGTCACAGGATGATGAAATGCC	196
Db		121	tggagctgggacgtttggatgggaattaccctgtgccatctcgacaggggtccaggatgaaatgcc	180
OY		197	TGCCTTGATGTCAAGCTGTGAACAACCAACAGAGAGACTGTATTGTGGTCTGGGAGAGATTG	256
Db		181	tgctttagatgtcaagtgtgaacaaccaaacagagagactgtattgttggtctgggagaattg	240
OY		257	AACCATTCCTTCCACAACTGCTGCATGTCCCTGTGGSTGAACAGAACAATGCTGCCTC	316
Db		241	aatacatcttcceacaactcgtgatgtccctgtgggtggaacagaacaatgcctgcctc	300
OY		317	CTGGCCAGACAGGACTGGGTAGTCCCAAAGAAATGGCCAATGACAGAGTGGCCAGCGCTC	376
Db		301	ctcgccagcacgagactgggtatgtccc aaagaatcgccaatlgagagatggtatagaaggcttc	360
OY		377	CTGGTGTGTGTTG-----CTGACCCCTGGACAAGACTAA	409
Db		361	tltagcgcagttgttcacagacctggtgatcttgtaatccagtgacctataaaggctcaga	420
OY		410	ACACTGACGGGATTCATCCTTGAGAGAG--AAGAGATGCTGTGCGCCTTTGAGACTCAC	467
Db		421	aaactacagggatgaattcttcaaatagagccgatgatctgtggtctttgactcat	480
OY		468	CAAGGCTTGCTTATTATAATTTGCTGTTAGTTTTGGAAATTTCTGTACAATT---AAC	524
Db		481	caaaagcttggtttag--catttgcagtttatctttaagaattctctgtgattaagaag	538
OY		525	ATAATTGTTAAAAATGGCCTTTCCTACCTGTGGTGTGTGT-----GTGATACGA	576
Db		539	ataatttattaaaggigtgtcctctaccctctgtgtgtgtgtgtgcgcgaacacgcttaga	588
OY		577	ATGCATGGAAGACGCGACACACAGAAAATGATCTTTGTTTATCTGTGACCGACACTGGA	636
Db		599	agtgtcataaaaaaggaagaagctcccaattgtaacccttaattaccatctctat	658
OY		637	ACATTGTTCCACAGAAAGACATGTTTGGTTTATCTGTGAGGGTTAAAAAATATAGATA	696
Db		659	acaacacaggcagtggaagcagttctgagacttcttcgatctctatgtgtgtatcsqttaaa	718
OY		697	ACGAATGTTACAGTAACAATAATAAATCATTTGAAAA	732
Db		719	aagaatgttacagtaacaataaagtcagtttaa	754
RESULT		10		
AAx87324		ID		
AAx87324		standard; cDNA; 754 BP.		
AAx87324;				
XX AC				
XX DT		27-SEP-1999 (first entry)		
XX DX				
XX XX		Human sensitive to apoptosis (SAG) gene mutant MM8.		
KW KW		SAG gene; sensitive to apoptosis; human; cancer; tumour;		
KM KM		neurodegenerative disease; muscular dystrophy; wound healing;		
KS KS		vulnerable; therapy; mutant; ds.		
OS OS		Homo sapiens.		
OS OS		Synthetic.		
XX FT				
XX FH		Key		
FT CDS		Location/Qualifiers		
FT FT		1..342		
FT FT		/+tag= a		
FT FT		mutation replace(262,T)		
XX FT		/+tag= b		
XX PN		/note= "C88S mutation"		
XX DN		MO9932514-A2.		
XX PD		01-JUL-1999		

[illegible]

XX MO9932514-A2.
 PN 01-JUL-1999.
 XX 15-DEC-1998; 98WO-US26705.
 XX 11-SEP-1998; 98US-0099840.
 PR 19-DEC-1997; 97US-0068179.
 XX (WARN) WARNER LAMBERT CO.
 PA Sun Y;
 PI WPI: 1999-430152/36.
 DR P-PSDB: AAY06506.
 XX SAG: Sensitive to Apoptosis Gene and related proteins, useful for
 PT promoting cell growth and protecting cells against apoptosis
 PS Claim 15; Page 77-78; 84pp; English.

XX This is the nucleotide sequence of human sensitive to apoptosis
 CC (SAG) mutant gene MM12, which codes for a SAG protein (see AAY06506),
 CC in which the Cys residue at position 73 of the native protein (see
 CC AAY06492) is replaced by a Ser residue owing to a mutation of codon
 CC 73 from TGT to AGC obtained by site-directed mutagenesis of SAG
 CC cDNA. This residue is in a protease inhibitor motif of SAG. Single
 CC and double SAG mutants (see AX87317-31) were made to determine
 CC the role of each cysteine residue of SAG in haem binding and SAG
 CC oligomerization. These properties were unaffected by the MM12
 CC mutation. SAG is a novel zinc finger protein that promotes
 CC cell growth, protects cells from apoptosis, scavenges oxygen
 CC radicals and can be used for the reversion of a tumour phenotype.
 CC SAG genes, and mutant SAG genes, can be used to protect cells from
 CC apoptosis induced by redox reagents. They can also be used for the
 CC recombinant production of SAG proteins, which are molecular targets
 CC in the development of drugs against neurodegenerative disorders,
 CC cancers and muscle dystrophy, and promoting wound healing.

XX Sequence 754 BP; 206 A; 155 C; 201 G; 192 T; 0 other;

Query Match 29.2%; Score 332.4; DB 20; Length 754;
 Best Local Similarity 70.5%; Pred. No. 4,7e-85;
 Matches 533; Conservative 0; Mismatches 181; Indels 42; Gaps 5;
 QY 17 ATGCCGAGCGTGGAGAGCGGCGAAGACCTGCTCTTCTTCGACATCCGGAGACGCA 76
 DB 1 atgcccgaagctgaaagaagcgaagaaacctgcgcctcctcctcgcggaagctca 60
 QY 77 GGCCTCAAGTCGGAGGCGCAGCAAGATGTTCTCTCAAGAAGTGGAAACCGGTAGCCATG 136
 DB 61 ggtccaagtcggaagcgaacaaagatgtctccctcaaaagaaagcgcggtggaacatg 120
 QY 137 TGAGAGCGGAGCGTGGTGCATACCTGTCATCTGACAGGTCAGAGTGAATGATGCC 196
 DB 121 tgaagctgggaagctgaaagtcgtaagtcgcacatctgcaggtccagtgatgaagtcgc 180
 QY 197 TGCCCTCGATGTAAGCTGAAAACAAGCAAGAGACTGTGTTGCTTGGGGAGAGCTG 256
 DB 181 tgccttagatgtcaagctgaaacaaagagacagtgctggtctggtggaagaaatgt 240
 QY 257 AACCATTCCTTCACAACTGCTGATGTCCTGTTGGTGAACAGAAATCGCTCCCT 316
 DB 241 aatcatctccctcacacatctgcatagtccctgtggtggaacaaagacatctgcctcct 300
 QY 317 CTGTGCGAGAGAGCTGGTAGTCCAAAGAAATGCGCAAAATGAGAGGTGGCCAGCGCTC 376
 DB 301 ctctgcagcagagactgtggtggtccaaagatcgcgaatgagagtggttaagaagcttc 360
 QY 377 CTGGTGTGTTG-----CTGACCTTGAGCAAAAGACTAA 409

DB 361 tttagcagtgcttcagagccctgtgagatctgttaatccagtgccctcaaaagctaga 420
 QY 410 ACATGCGAGGATTCATCTTGAGAGAG--AGAGATGCTGTGGCCCTTGAGACTCAC 467
 DB 421 acactacagggatgaattcttcaaataggaagccgatgagatctgtggtgactcat 480
 QY 468 CAAGGCTGCTTTATTAATTTGTTGTTAGTTTGGGAAATTCCTACAAATT--AAG 524
 DB 481 caaagcctgtgtag--cattgtcagtttattcttcaagaattctctgtgtaagaag 538
 QY 525 ATAAATTTGTTAAATGAGCCTTCTTACTCTGCTGTGTGT-----GTGATACGA 576
 DB 539 atattattaaagtggtccttccactctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 598
 QY 577 ATGATAGAAAGACGAGAACACAGAAATGATCTTTGTTTATCTGTACCAAGACTGGA 636
 DB 599 agtgcataaaagaagaagagctccaatgaaatcattatccatttccattctat 658
 QY 637 ACATTTGTTTCACAGAAAGACATTTGTTGTTTATGCTTGAGGCTTAAATAATAGATA 696
 DB 659 acaacagcgctgtgaagcagtttcgagacttttcgagtcttatgtgtgatacgttaaa 718
 QY 697 ACGAATGTTACGTACATAAATAATGATTGAAA 732
 DB 719 aagaatgttacagtaacaataaagtcagtttaa 754

RESULT 14
 ID AAX87331 standard; cDNA; 754 BP.
 AC AAX87331;
 AC 27-SEP-1999 (first entry)
 DT Human sensitive to apoptosis (SAG) gene mutant MM15.
 DE SAG gene; sensitive to apoptosis; human; cancer; tumour;
 KW neurodegenerative disease; muscular dystrophy; wound healing;
 KW vulnereary; therapy; mutant; ds.
 XX Homo sapiens.
 OS Synthetic.
 OS Key Location/Qualifiers
 FH CDS 1..342
 FT /*tag= a
 FT mutation replace(139,C)
 FT /*tag= b
 FT /*note= "C47S mutation"
 W09932514-A2.
 PN 01-JUL-1999.
 XX 15-DEC-1998; 98WO-US26705.
 XX 11-SEP-1998; 98US-0099840.
 PR 19-DEC-1997; 97US-0068179.
 XX (WARN) WARNER LAMBERT CO.
 PA Sun Y;
 PI WPI: 1999-430152/36.
 DR P-PSDB: AAY06509.
 PT SAG: Sensitive to Apoptosis Gene and related proteins, useful for
 PT promoting cell growth and protecting cells against apoptosis
 PS Claim 15; Page 82-83; 84pp; English.
 XX This is the nucleotide sequence of human sensitive to apoptosis

CC (SAG) mutant gene MM15, which codes for a SAG protein (see AAY05609)
CC in which the Cys residue at position 47 of the native protein (see
CC AAY06492) is replaced by a Ser residue owing to a mutation of codon
CC 47 from TGC to AGC obtained by site-directed mutagenesis of SAG
CC cDNA. This residue is in a GADPH binding site of SAG. Single
CC and double SAG mutants (see AAX87317-31) were made to determine
CC the role of each cysteine residue of SAG in haem binding and SAG
CC oligomerization. These properties were unaffected by the MM15
CC mutation. SAG is a novel zinc finger protein that promotes
CC cell growth, protects cells from apoptosis, scavenges oxygen
CC radicals and can be used for the reversion of a tumour phenotype.
CC SAG genes, and mutant SAG genes, can be used to protect cells from
CC apoptosis induced by redox reagents. They can also be used for the
CC recombinant production of SAG proteins, which are molecular targets
CC in the development of drugs against neurodegenerative disorders,
CC cancers and muscle dystrophy, and promoting wound healing.

SQ Sequence 754 BP; 206 A; 155 C; 201 G; 192 T; 0 other;

Query Match	29.28;	Score 332.4;	DB 20;	Length 754;
-------------	--------	--------------	--------	-------------

Matches 533; Conservative 0; Mismatches 181; Indels 42; Gaps 5;

Matches 533; Conservative 0; Mismatches 181; Indels 42; Gaps 5;

OY	17	ATGGCCGACGCGGAGGACGGGAGACACCCGGGTCTCTTCGCACTCCGGACGCA	76
Db	1	atggccgacgcgaggagagagaaacctcgcccttgccctcaccctcggagactaa	60
OY	77	GGCTCCAAAGTCGGGAGGCGCAAGATGTTCTCTCAAGAAAGTGAACGGCGTAGCCATG	136
Db	61	ggctccaaagtcgggagggagaaagatgtcttcctccaagaagcggaaacggcggccaa	120
OY	137	TGGAAGCTGGGACGTTGAGTGGATACCTGTGCCATCTGCAGGGCTCAGGTGATGATGCC	196
Db	121	tggaagctgggacgttgagtgagtcgaacacgtgcacatctgcagggtccagggtga	180
OY	197	TGCCCTGCATTCATCAAGCTGAAACAAACAGAGAGACTGTTGGTGGCTGGGAGAGTGT	256
Db	181	tgccctgcattcatcaagctgaaacaaacaaagaagactgtgtgtgtctgtggagaaat	240
OY	257	AACCATTCCTCCCAACACTGCTGCATGTCCCTGTGGGTGAACAGAACATCCCTCCCT	316
Db	241	aaccattcctccccaacactgctgatagtccctgtgggtgaacaaagaaacaaatcgct	300
OY	317	CTGTGCCAGCAGGACTGGGTATGTCCAAAGAAATGGCAATGAGAGTGGCCAGCGCTC	376
Db	301	ctgtgccagcagagactggtgtgtccaaagaatcggcaaatgagagatggtlaagaagctc	360
OY	377	CTGTGTGTGTTG-----CTGACCCCTGGGCAAAAGCTAA	409
Db	361	ctgtgtgtgttg-----ctgacctgggcaaaagctaa	420
OY	410	ACACTGCAGGGGATTCATCCTTGAGAGAG--AGAGATGCTGTGGCGCTTGAGACTCAC	467
Db	421	acaactcagggagatgaactcttcaataaagaagccgatcgtgtgtctcttgcactac	480
OY	468	CAAAAGCCTTGCTTATTAATTATTTGCTGTGTAGTTTGGGAAATTCCTTACAAATT---	524
Db	481	caaaagccttgtag--cattgtgcagtttatcttcgaagaattctctgtgattaagaag	538
OY	535	ATTAATTGTTAAAAATGGCCCTTCCCTCACTCTGTGGTGTGTG-----GTGATACGA	576
Db	539	ataatttatataaagtgtgtccctctcactctctgtgtgtgtgtgcgcacaacagcttag	598
OY	577	ATGCATGAAAGAGGAGAGAACCCAGAAATATGATTTTGTTATCTGTACCCACGACTGGA	636
Db	599	atgtcctataaaaagaagaagcgtcccaatgtaacactcctataatlaaccacttcat	658
OY	637	ACATTGTGTTACAGCAAGAAACATTGTTGTGTTATATCTTGAGGGTTAAAAATAGTAA	696
Db	659	acaaagcgcagtgaaagcagtttcggaacatttctgcagtcttatgtgtatcagttaaaa	718
OY	697	ACGAATGTTACAGTACAAATAAATGATCATTTGAATA	732

Db 719 aagaatgttacgtacaacaataaagtcagtttaa 754

RESULT 15
AAx87319
ID AAx87319 standard; cDNA; 754 BP.

AC AAX87319;

DT 27-SEP-1999 (first entry)

DE Human sensitive to apoptosis (SAG) gene mutant MM3.

KW SAG gene; sensitive to apoptosis; human; cancer; tumour;

KW vulnerary; therapy; mutant; ds.

OS Homo sapiens.

XX

FT CDS

FT mutation

ET

FT

X

XX
XX
2
1
4
4
4
2
2
2

XX
DE 15-DEC-1960XX 11-SEP-1966
PBPR 19-DEC-1955
yyPA (WARN) WPA
YY

PI Sun Y;
YY

DR WPI; 1999-
D-DEPB; A2

XX
DT Const

PT promoting
xy

PS Claim 15;
yy

CC This is the
CC (EAC) mut-

in which the

CC mutations
CC site-directed

CC binding si
CC were made

CC of SAG in
CC reduced by

CC novel zinc
CC from apert

CC reversion

CC reagents.

CC drugs against

XXXXXXXXXXXX

Sequence 754 BP; 207 A; 155 C; 201 G; 191 T; 0 other;

Query Match	29.0%;	Score 330.8;	DB 20;	Length 754;
Best Local Similarity	70.4%;	Pred. No. 1.3e-84;		
Matches 532;	Conservative 0;	Mismatches 182;	Indels 42;	Gaps. 5

QY	17	ATGGCGCAGCGTGGAGGAGCGGCGAGGAAACCCCTGGTCTCTTTCCTTGCGACTCCGGGAGCGGA	76
Db	1	atggcgcaagctggaaagcggagaaacctctgcccctgcctctcaaccgggagctca	60
QY	77	GGCTCCAAAGTCGGGAGGCGCAGAGATGTTCTCTCTCAAGAAGTGAACGCGGTATGCCATG	136
Db	61	ggctccaagctcgaggagcgacaagatgctctccctcaagaagtggaaacggctgycgcctg	120
QY	137	TGGAGCTGGGACCTTGATGTCGATCTGTCATCTGCGATCTGCAAGGGTCCAGGATGATGATGCC	196
Db	121	tggagctgggacgcttgatgtcgaatcgtgacacgaagcgccatcagcagggtccaaagtgatgtgc	180
QY	197	TGCGCTTGATGTCGAAGCTGGAACCAAGCAGAGACTGTGTTGTGGTCTGGGAGAGTGT	256
Db	181	tgccttagatgtcgaagctggaacaacaacagagactgtgtgtgtcttggtagaagtgt	240
QY	257	AACCATTCCTTCCACACACTGCTGATGTCCTCCCTGGGTGAACAAGAACAAATGCGTCCCT	316
Db	241	aatacttcctcccaactcctgtgatagtcccttggttgatgaacaagaacatcgtccct	300
QY	317	CTTGCCCGACGAGCATGGGTGTGTCCAAGAAGATGGCAAAATGAGAGGTGGCCCGGCGCTC	376
Db	301	ctctgcgcagcagagcctgtgtgtctcaagaatcgcgaatggagatggtttagaagcttc	360
QY	377	CTGGTGTGGTGG-----CTGACCCGTGGACCAAGACTAA	409
Db	361	cttagcgagctgtgtccagagccctgtgtgatactcagtgcccttaacaaggtctaga	420
QY	410	ACACTGACGGGATTCATCCTTGAAGAG--AGAGATGCTGTGCGCCTTGAAGCTCAC	467
Db	421	aactacacagggatgtaattcttcaataatagagccgagatgtgtgtcttcttgactat	480
QY	468	CAAAAGCTTCGCTTATTAATATGTCGTTTAAGTTTGGGAAATTCCTACAAAT---AAG	524
Db	481	caaaagctctgttag--cattgtcagatttlaattcctaagaattctctgtgataaaga	538
QY	525	ATAATTTGTTAAAAATGCGCTTCTCTCACTCTGGTGTGTGTGT-----GTGATACGA	576
Db	539	ataatttataaagtggtctctctcctcactcctctgtgtgtgtgtgcgcacacagcttag	598
QY	577	ATGCTATAAAGACGAGAACACCGAAAAATGATCTTTGTTATCTGTACCCAGACTGGA	636
Db	599	atgtctataaagaagaagagcctccaatgtgaatcactataattaccacttctat	658
QY	637	ACATTGTTTTCACAGAAACATTTGTTGTTTATGCTTGAAGGTTAAAAAATTAGATAA	696
Db	659	acaacacgacgtggaagcagttctcgagacttttcgagtcttaatgtgtatcagttaa	718
QY	697	ACGAATGTTACGTAACAATAATGATGTGAATA 732	
Db	719	aagaaatgttaccagttacaataaagtgaagtttaa 754	

Search completed: July 25, 2002, 20:25:44
Job time: 10967 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 25, 2002, 15:11:49 ; Search time 102.69 Seconds

(without alignments)
2726.869 Million cell updates/sec

Title: US-09-509-779-1

Perfect score: 1140

Sequence: 1 GTCTGCGCGCGCGCCATG.....ACTATTCATCATTAATG 1140

Scoring table: IDENTITY_NUC

Gapop 10.0 ; Gapext 1.0

Searched: 38353 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued_Patents_NA :
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq : *
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq : *
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq : *
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq : *
5: /cgn2_6/ptodata/2/ina/PTUS.COMB.seq : *
6: /cgn2_6/ptodata/2/ina/Backfiles1.seq : *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	3.5	549	1	US-07-991-867B-28
2	40	3.5	549	1	US-08-107-755A-28
3	40	3.5	549	2	US-08-544-332-28
4	40	3.5	1511	1	US-07-991-867B-8
5	40	3.5	1511	1	US-08-107-755A-8
6	40	3.5	1511	2	US-08-544-332-8
7	40	3.5	4810	3	US-08-852-629-11
8	40	3.5	4838	3	US-08-852-629-15
9	38.2	3.4	7218	1	US-08-232-463-14
10	36.4	3.2	1037	2	US-08-824-405-1
11	36	3.2	2861	1	US-08-299-953-1
12	36	3.2	2861	1	US-08-459-415-1
13	36	3.2	2861	4	US-09-066-687-1
14	36	3.2	2861	5	PCT-US95-11231-1
15	36	3.2	3881	1	US-08-299-953-2
16	36	3.2	3881	1	US-08-459-415-2
17	36	3.2	3881	4	US-09-066-687-2
18	36	3.2	3881	5	PCT-US95-11231-2
19	35.8	3.1	3019	4	US-09-359-161-2
20	35.6	3.1	3701	4	US-08-845-258-10
21	35.6	3.1	3701	4	US-08-990-571-10
22	35.6	3.1	3701	4	US-08-723-142A-10
23	35.4	3.1	4818	3	US-08-817-926-27
24	35	3.1	3181	1	US-08-655-086-1
25	34.8	3.1	767	4	US-08-998-416-472
26	34.6	3.0	591	2	US-08-943-208-1
27	34.6	3.0	1982	3	US-08-747-221B-13

28	34.6	3.0	1982	3	US-08-747-221B-15	Sequence 15, Appl
29	34.6	3.0	1982	4	US-09-005-051-13	Sequence 13, Appl
30	34.6	3.0	1982	4	US-09-005-051-15	Sequence 15, Appl
31	34.6	3.0	2144	3	US-08-747-221B-57	Sequence 57, Appl
32	34.6	3.0	2144	3	US-08-747-221B-59	Sequence 59, Appl
33	34.6	3.0	2144	4	US-09-005-051-57	Sequence 57, Appl
34	34.6	3.0	2144	4	US-09-005-051-59	Sequence 59, Appl
35	34.4	3.0	285	1	US-08-435-040-1	Sequence 1, Appl
36	34.4	3.0	285	4	US-09-020-216-1	Sequence 1, Appl
37	34.4	3.0	7904	1	US-08-316-238B-1	Sequence 1, Appl
38	34.4	3.0	7904	1	US-08-316-238B-2	Sequence 2, Appl
39	34.4	3.0	7904	1	US-08-410-005-1	Sequence 1, Appl
40	34.4	3.0	7904	3	US-08-929-140-1	Sequence 1, Appl
41	34.4	3.0	7904	4	US-09-560-579A-1	Sequence 1, Appl
42	34.2	3.0	2555	2	US-08-693-457-3	Sequence 3, Appl
43	34.2	3.0	2555	4	US-09-265-731-3	Sequence 3, Appl
44	34	3.0	3435	4	US-09-329-685A-1	Sequence 1, Appl
45	33.8	3.0	1361	4	US-09-232-191-20	Sequence 20, Appl

ALIGNMENTS

RESULT 1
US-07-991-867B-28
; Sequence 28, Application US/07991867B
; Patent No. 5476781
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Hall, Richard L.
; APPLICANT: Gruidl, Michael E.
; TITLE OF INVENTION: No. 5476781el Entomopoxvirus Expression System
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/991,867B
; FILING DATE: 12-DEC-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: NO 92/14818
; FILING DATE: 12-FEB-1992
; APPLICATION NUMBER: US 07/827,685
; FILING DATE: 30-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/657,584
; FILING DATE: 19-FEB-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: UFI14.C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 549 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; US-07-991-867B-28

1


```

Oy 941 ACACCTAGTTCGTAAGAAGACTCTTTCTGTTTTGCGGAGGTGTGATGATGTTAG 1060
Db 220 AACCATACCTTTTTTTTTTATTATTACCAATTTTATCACAATAATTTCTAAATCATTTTC 279
Oy 1001 TCAGAAAATATTAGTACGAAAATGCGTTTACTAGTATPAACACTGAAATTCATTATGCAATGT 1060
Db 280 TTCAAAAATATTGACACATCATCTATGCGCAATPAATATCATATATATCTACGATATGATTTTC 339
Oy 1061 TTTATTAATAATATGTGCTTTGAGTTATTAAGATTGATATATCTCT 1108
Db 340 ATTAATAATAATATTGTTTAAATGTAATAAATATCTTATTTAATAT 387

RESULT 4
US-07-991-867B-8
; Sequence 8, Application US/07991867B
; Patent No. 5476781
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Hall, Richard L.
; APPLICANT: Gruidl, Michael E.
; TITLE OF INVENTION: No. 54/6781el Entomopoxvirus Expression System
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/991,867B
; FILING DATE: 12-DEC-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO 92/14818
; FILING DATE: 12-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/827,685
; FILING DATE: 30-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/657,584
; FILING DATE: 19-FEB-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: UFI114.C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ. ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1511 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Amsacta moorei entemopoxvirus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: complement (18..218)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: complement (234..782)
; FEATURE:
; NAME/KEY: CDS

```

```

1      LOCATION: 852..1511
2      US-07-991-867B-B
3
4      Query Match          3.5%; Score 40; DB 1; Length 1511;
5      Best Local Similarity 52.4%; Pred. No. 0.065;
6      Matches 88; Conservative 0; Mismatches 80; Indels 0; Gaps 0
7
8      Oy 941 ACCACTGAGTCTAAAGACGTCTTTTCTGTTTGGCCGAGGTGCTATGTGATGTTTGG 1000
9          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
10     Db 453 AACCATACACTTTTATTTTATTATTTAGCCATTTATTCACAAAAATGTCTAAATTCATTTTC 512
11
12     Oy 1001 TCAAAATATTTGCTGGAAGAAATGGCTTACTAGTATACACGTGACATGCTATTCGCAATGT 1060
13         | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
14     Db 513 TTCAAAAATATGACACTCATCTATTCATCCATATATATCATATATATATGACATATTTGATTTTC 572
15
16     Oy 1061 TTTAATTAATATATCTGCTTTGAGTTATATTAAGTTGATATATATCTCT 1108
17         | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
18     Db 573 ATTAATTAATATATTTGTTTAAATGATATATATCTTATTTAATAT 620
19
20 RESULT 5
21 US-08-107-755A-8
22 Sequence 8, Application US/08107755A
23 Patent No. 5721352
24
25 GENERAL INFORMATION:
26 APPLICANT: Moyer, Richard W.
27 APPLICANT: Hall, Richard L.
28 APPLICANT: Gruidl, Michael E.
29 TITLE OF INVENTION: No. 5721352e1 Entomopoxvirus Expression System
30 NUMBER OF SEQUENCES: 40
31 CORRESPONDENCE ADDRESS:
32 ADDRESSEE: David R. Saliwanchik
33 STREET: 2421 N.W. 41st Street, Suite A-1
34 CITY: Gainesville
35 STATE: Florida
36 COUNTRY: U.S.A.
37 ZIP: 32606
38
39 COMPUTER READABLE FORM:
40 MEDIUM TYPE: Floppy disk
41 COMPUTER: IBM PC compatible
42 OPERATING SYSTEM: PC-DOS/MS-DOS
43 SOFTWARE: Patentin Release #1.0, Version #1.25
44 CURRENT APPLICATION DATA:
45 APPLICATION NUMBER: US/08/107,755A
46 FILING DATE: 19-AUG-1993
47 CLASSIFICATION: 435
48
49 PRIOR APPLICATION DATA:
50 APPLICATION NUMBER: US 07/827,658
51 FILING DATE: 30-JAN-1992
52 PRIOR APPLICATION DATA:
53 APPLICATION NUMBER: US 07/657,584
54 FILING DATE: 19-FEB-1991
55 ATTORNEY/AGENT INFORMATION:
56 NAME: Saliwanchik, David R.
57 REGISTRATION NUMBER: 31,794
58 REFERENCE/DOCKET NUMBER: 0F114.C2
59 TELECOMMUNICATION INFORMATION:
60 TELEPHONE: (904) 375-8100
61 TELEFAX: (904) 372-5800
62 INFORMATION FOR SEQ ID NO: 8:
63 SEQUENCE CHARACTERISTICS:
64 LENGTH: 1511 base pairs
65 TYPE: nucleic acid
66 STRANDEDNESS: double
67 TOPOLOGY: unknown
68 MOLECULE TYPE: DNA (genomic)
69 ORIGINAL SOURCE:
70 ORGANISM: Amsacta moorei entomopoxvirus
71 FEATURE:
72 NAME/KEY: CDS
73 LOCATION: complement (18..218)
74 FEATURE:

```

NAME/KEY: CDS
LOCATION: complement (234..782)
FEATURE:
NAME/KEY: CDS
LOCATION: 852..1511
US-08-107-755A-8

Query Match 3.5%; Score 40; DB 1; Length 1511;
Best Local Similarity 52.4%; Pred. No. 0.065;
Matches 88; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 941 ACCACTAGTCTTAAGAAGCTGTTTCTGTTTGGCGAAGCTGTATGTTAG 1000
DB 453 AACCAATACCTTTTATTTATTTACCAATTTTATCACAATTTGTTAAATCATTTTC 512
QY 1001 TCAAAATATTAGTGAAGAAATGGCTTACTAGTATACAGCTGATGATTCATTCAGATGT 1060
DB 513 TTCAAAAATTTGACACTCTATGCGCAATTAATCATATATATCATGATATGATTTTC 572
QY 1061 TTTAATAAATATTGTGCTTGTGATTTAAAGTTGATATATCTCT 1108
DB 573 ATTAAATTAATTTTGTTTTATGTAATTAATATTTCTTTATTTAAT 620

RESULT 6

US-08-544-332-8
Sequence 8, Application US/08544332
Patent No. 5935777
GENERAL INFORMATION:
APPLICANT: Moyer, Richard W.
APPLICANT: Hall, Richard L.
APPLICANT: Gruidl, Michael E.
TITLE OF INVENTION: No. 5935777e1 Entomopoxvirus Expression System
NUMBER OF SEQUENCES: 77
CORRESPONDENCE ADDRESS:
ADDRESSEE: Gerard H. Bencen
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/544,332
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/991,867
FILING DATE: 07-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/107,755
FILING DATE: 19-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO 92/14818
FILING DATE: 12-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,685
FILING DATE: 30-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/657,584
FILING DATE: 19-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: Bencen, Gerard H.
REGISTRATION NUMBER: 35,746
REFERENCE/DOCKET NUMBER: UF114.C4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800

INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1511 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE: Amsacta moorei entemopoxvirus
ORGANISM: Amsacta moorei entemopoxvirus
FEATURE:
NAME/KEY: CDS
LOCATION: complement (18..218)
FEATURE:
NAME/KEY: CDS
LOCATION: complement (234..782)
FEATURE:
NAME/KEY: CDS
LOCATION: 852..1511
US-08-544-332-8

Query Match 3.5%; Score 40; DB 2; Length 1511;
Best Local Similarity 52.4%; Pred. No. 0.065;
Matches 88; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 941 ACCACTAGTCTTAAGAAGCTGTTTCTGTTTGGCGAAGCTGTATGTTAG 1000
DB 453 AACCAATACCTTTTATTTATTTAGCCATTTTATCACAATTTGTTCAATCATTTTC 512
QY 1001 TCAAAATATTAGTGAAGAAATGGCTTACTAGTATACAGCTGATGATTCATTCAGATGT 1060
DB 513 TTCAAAAATTTGACACTCTATGCGCAATTAATCATATATATCATGATATGATTTTC 572
QY 1061 TTTAATAAATATTGTGCTTGTGATTTAAAGTTGATATATCTCT 1108
DB 573 ATTAAATTAATTTTGTTTTATGTAATTAATATTTCTTTATTTAAT 620

RESULT 7

US-08-852-629-11
Sequence 11, Application US/08852629
Patent No. 6106825
GENERAL INFORMATION:
APPLICANT: Moyer, Richard W.
APPLICANT: Li, Yi
APPLICANT: Hall, Richard L.
TITLE OF INVENTION: ENTOMPOXYVIRUS-VERTEBRATE GENE DELIVERY
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: U.S.A.
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/852,629
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Bencen, Gerard H.
REGISTRATION NUMBER: 35,746
REFERENCE/DOCKET NUMBER: UF-184
TELECOMMUNICATION INFORMATION:
TELEPHONE: 407-426-7500
TELEFAX: 407-839-8589
INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:
LENGTH: 4810 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-852-629-11

Query Match 3.5%; Score 40; DB 3; Length 4810;
Best Local Similarity 52.4%; Pred. No. 0.12;
Matches 88; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 941 ACCACTAGTCTCAAGAACTGTTTCTGTTTGGCGAAGTGTGATGTTAG 1000
DB 459 AACCAATACCTTTTATTTATAGCAATTTTATCACAATAATGTTCTAAATCATTTTC 518

QY 1001 TCATAATATTTAGTGAAGAAATGCTTACTAGTATTAACCTGAAGTTCAATATGCAATGT 1060
DB 519 TTCAAAAATTTGACACTCTCTATGCCAATATATCATATTTATCTACCATATTTGATTTC 578

QY 1061 TTTAATAAATATTTGCTTGTGATTTAATACTTGATATATATCTCT 1108
DB 579 ATTAATTAATATTTGTTTATGTATTAATAATATCTTTATTTAATAT 626

RESULT 8

US-08-852-629-15
Sequence 15, Application US/08852629
Patent No. 6106825

GENERAL INFORMATION:

APPLICANT: Moyer, Richard W
APPLICANT: Li, Yi
TITLE OF INVENTION: ENTOMOPOXYVIRUS-VERTEBRATE GENE DELIVERY
TITLE OF INVENTION: VECTOR AND METHOD
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: U.S.A.
ZIP: 32606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/852,629
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Bence, Gerard H
REGISTRATION NUMBER: 35,746
REFERENCE/DOCKET NUMBER: UF-184
TELECOMMUNICATION INFORMATION:
TELEPHONE: 407-426-7500
TELEFAX: 407-839-8589
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 4838 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-852-629-15

Query Match 3.5%; Score 40; DB 3; Length 4838;
Best Local Similarity 52.4%; Pred. No. 0.12;
Matches 88; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 941 ACCACTAGTCTCAAGAACTGTTTCTGTTTGGCGAAGTGTGATGTTAG 1000
DB 459 AACCAATACCTTTTATTTATAGCAATTTTATCACAATAATGTTCTAAATCATTTTC 518

QY 1001 TCATAATATTTAGTGAAGAAATGCTTACTAGTATTAACCTGAAGTTCAATATGCAATGT 1060
DB 519 TTCAAAAATTTGACACTCTCTATGCCAATATATCATATTTATCTACCATATTTGATTTC 578

QY 1061 TTTAATAAATATTTGCTTGTGATTTAATACTTGATATATATCTCT 1108
DB 579 ATTAATTAATATTTGTTTATGTATTAATAATATCTTTATTTAATAT 626

RESULT 9

US-08-232-463-14/C
Sequence 14, Application US/08232463
Patent No. 5670367

GENERAL INFORMATION:

APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOMLOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZ9pt-F15
US-08-232-463-14

Query Match 3.4%; Score 38.2; DB 1; Length 7218;
Best Local Similarity 6.6%; Pred. No. 0.48;
Matches 28; Conservative 207; Mismatches 190; Indels 0; Gaps 0;

QY 75 CAGGCTCAAGTGGAGGCGACAGATGTTCTCTCAAGAGTGAACGCGTAGCCA 134

LENGTH: 1037 base pairs

TOPOLOGY: linear

RESULT 11
 US-08-299-953-1
 ; Sequence 1, Application US/08299953
 ; Patent No. 5646333
 ; GENERAL INFORMATION:
 APPLICANT: Dobres, Michael S. and Mandaci, Sevaur
 TITLE OF INVENTION: A plant promoter useful for directing the
 TITLE OF INVENTION: Expression of Foreign Proteins to the plant epidermis
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESS: Woodcock Washburn Kurtz Mackiewicz & No. 5646333rirs
 STREET: One Liberty Place 46th. Floor
 CITY: Philadelphia
 STATE: PA
 ZIP: 19103
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/299,953
 FILING DATE: Herewith
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Beardell, Lori Y.
 REGISTRATION NUMBER: 34,293
 REFERENCE/DOCKET NUMBER: NOVA-0003
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 215-564-8960
 TELEFAX: 215-568-3439
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2861 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)

DB 1070 GGTGTTTGAAGAAAAATATATCATATATAATCAATTTTACAGTTTAAAAAGAAAT 1129
 OY 1064 AATAAATATATGCTTGTAGTTATTAAGTTGATATATCTTTAAATCAATTAAC 1123
 DB 1130 AATATTAGTTTATTTATTAATCACTAGAAATTTATATATCTTTTATTAATATA 1189
 OY 1124 AATTCATCAATTAAT 1139
 DB 1190 AATTATCTCTCCAT 1205

RESULT 14

PCT-US95-11231-1
 ; Sequence 1, Application PC/US9511231
 ; GENERAL INFORMATION:
 ; APPLICANT: Dobres, Michael S. and Mandaci, Sevnuur
 ; TITLE OF INVENTION: A Plant Promoter Useful for Directing the Expression
 ; TITLE OF INVENTION: of Foreign Proteins to the Plant Epidermis
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & Norris
 ; STREET: One Liberty Place 46th. Floor
 ; CITY: Philadelphia
 ; STATE: PA
 ; ZIP: 19103
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US95/11231
 ; FILING DATE: Herewith
 ; CLASSIFICATION:
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: 08/299,953
 ; FILING DATE: September 2, 1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Beardell, Lori Y.
 ; REGISTRATION NUMBER: 34,293
 ; REFERENCE/DOCKET NUMBER: NOVA-0016
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 215-564-8960
 ; TELEFAX: 215-568-3439
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2861 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; HYPOTHEICAL: NO
 ; ANTI-SENSE: NO
 ; PCT-US95-11231-1

Query Match 3.2%; Score 36; DB 5; Length 2861;
 Best Local Similarity 49.0%; Pred. No. 1.3; Mismatches 100; Indels 0; Gaps 0;
 Matches 96; Conservative 0;

OY 944 ACTTAGTCTAAGAACTGTTTCTGTTTGGCGAAGTTGATTTGATGTTTAACTCA 1003
 DB 1010 ACTTGTCTTTATTAATATTTCTGACATATGATTAATCTCCCTATTCATTTTATTA 1069
 OY 1004 AATAATATTAGTAGAATATGCTTACTAGATTAACCTGAAGTTCATTATGCAATGTTT 1063
 DB 1070 GGTGTTTGAAGAAAAATATATCAATTAATCAATTTTACAGTTTAAAAAGAAAT 1129
 OY 1064 AATAAATATATGCTTGTAGTTATTAAGTTGATATATCTTTAAATCAATTAAC 1123
 DB 1130 AATATTAGTTTATTTATTAATCACTAGAAATTTATATATCTTTTATTAATATA 1189
 OY 1124 AATTCATCAATTAAT 1139

DB 1190 AATTATCTCTCCAT 1205

RESULT 15

US-08-299-953-2
 ; Sequence 2, Application US/08299953
 ; Patent No. 5646333
 ; GENERAL INFORMATION:
 ; APPLICANT: Dobres, Michael S. and Mandaci, Sevnuur
 ; TITLE OF INVENTION: A Plant Promoter Useful for Directing the
 ; TITLE OF INVENTION: Expression of Foreign Proteins to the Plant Epidermis
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5646333ris
 ; STREET: One Liberty Place 46th. Floor
 ; CITY: Philadelphia
 ; STATE: PA
 ; ZIP: 19103
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/299,953
 ; FILING DATE: Herewith
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Beardell, Lori Y.
 ; REGISTRATION NUMBER: 34,293
 ; REFERENCE/DOCKET NUMBER: NOVA-0003
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 215-564-8960
 ; TELEFAX: 215-568-3439
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3881 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; HYPOTHEICAL: NO
 ; ANTI-SENSE: NO
 ; US-08-299-953-2

Query Match 3.2%; Score 36; DB 1; Length 3881;
 Best Local Similarity 49.0%; Pred. No. 1.5; Mismatches 100; Indels 0; Gaps 0;
 Matches 96; Conservative 0;

OY 944 ACTTAGTCTAAGAACTGTTTCTGTTTGGCGAAGTTGATTTGATGTTTAACTCA 1003
 DB 1010 ACTTGTCTTTATTAATATTTCTGACATATGATTAATCTCCCTATTCATTTTATTA 1069
 OY 1004 AATAATATTAGTAGAATATGCTTACTAGATTAACCTGAAGTTCATTATGCAATGTTT 1063
 DB 1070 GGTGTTTGAAGAAAAATATATCAATTAATCAATTTTACAGTTTAAAAAGAAAT 1129
 OY 1064 AATAAATATATGCTTGTAGTTATTAAGTTGATATATCTTTAAATCAATTAAC 1123
 DB 1130 AATATTAGTTTATTTATTAATCACTAGAAATTTATATATCTTTTATTAATATA 1189
 OY 1124 AATTCATCAATTAAT 1139
 DB 1190 AATTATCTCTCCAT 1205

Search completed: July 25, 2002, 19:07:32
 Job time: 14143 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 25, 2002, 18:05:39 ; Search time 2540.26 Seconds
(without alignments)
4006.166 Million cell updates/sec

Title: US-09-509-779-3

Perfect score: 754

Sequence: 1 ATGCGCAGCAGTGAAGACGG.....CAATTAAGTGCACCTTAA 754

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estnu:*
5: em_estov:*
6: em_estopl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_hic:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	689	91.4	793	9	AL547435
2	664.8	88.2	949	10	AL547435
3	661.6	87.7	797	10	AL547435
4	659.8	87.5	918	10	AL547435
5	649	86.1	822	10	AL547435
6	647.4	85.9	818	10	AL547435
7	644.4	85.5	805	10	AL547435
8	643	85.3	856	10	AL547435
9	640.2	84.9	754	10	AL547435
10	634.4	84.1	724	10	AL547435
11	624.6	82.8	855	10	AL547435
12	623.6	82.7	893	10	AL547435
13	616.2	81.7	797	10	AL547435
14	609	80.8	856	10	AL547435
15	608.8	80.7	767	10	AL547435
16	607.4	80.6	702	9	AL547435
17	606.6	80.5	847	10	AL547435

18	606	80.4	748	10	BG576768
19	597.2	79.2	904	10	BG742338
20	594.2	78.8	716	10	BG761806
21	592.8	78.6	840	10	BG037022
22	589.8	78.2	827	10	B1828930
23	588.8	78.1	675	10	B1858784
24	587.2	77.9	947	10	BG111792
25	582.4	77.2	745	10	BG121625
26	581.8	77.2	734	10	BG397308
27	575.2	76.3	864	10	B1757881
28	574.8	76.2	793	10	BG716023
29	571.8	75.8	651	10	BG714665
30	555.4	73.7	595	10	BF033587
31	554.2	73.5	781	10	BG026092
32	553.6	73.4	670	9	AM001261
33	553.4	73.4	793	10	BG111145
34	546.2	72.4	715	10	BG718227
35	532.2	70.6	696	10	BG715867
36	530.2	70.3	600	10	B1668630
37	515.4	68.4	600	9	AA521294
38	509.2	67.5	673	10	BE398022
39	488.4	64.8	565	10	BG054821
40	486.4	64.5	596	10	BG054726
41	485.8	64.4	604	9	AA521231
42	479.4	63.6	598	10	BG181095
43	473.8	62.8	561	9	AT769519
44	473.4	62.8	549	9	AA583455
45	471.4	62.5	555	9	AA805121

ALIGNMENTS

RESULT 1
AL547435 793 bp mRNA linear EST 16-FEB-2001
LOCUS AL547435 LTI_NFL006.PL2 Homo sapiens cDNA clone CS0D1011YW21 5
DEFINITION AL547435 LTI_NFL006.PL2 Homo sapiens cDNA clone CS0D1011YW21 5
prime, mRNA sequence.

ACCESSION AL547435.1 GI:12881506
VERSION AL547435
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 793)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seq@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES

source
1..793
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1011YW21"
/clone_1b="LTI_NFL006.PL2"
/tissue_type="Placenta"
/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with NotI and cloned into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact: Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@life.com
http://fulllength.invitrogen.com"

BASE COUNT 230 a 163 c 203 g 195 t 2 others
ORIGIN

Db	91	GGCTCCAAAGTGGGAGGGCGACAAAGATGTTCTTCCTCAAGAAAGTGGAAAGCGGGTGGCCATG	150
QY	121	TGGAGCTGGGACGTTGGAGTGGCATGCTGGCCATCTGACAGGTCACAGTATGGATGCC	180
Db	151	TGGAGCTGGGACGTTGGAGTGGCATGCTGGCCATCTGACAGGTCACAGTATGGATGCC	210
QY	181	TGCTTTGATGATGTCAAAGCTGGAACAAACAAACAAAGGACCTGTGTTGGCTGGGGACAATGT	240
Db	211	TGCTTTGATGATGTCAAAGCTGGAACAAACAAACAAAGGACCTGTGTTGGCTGGGGACAATGT	270
QY	241	AATCATTTCTTCACCAACTGCTGCATGTCCCTGTGGGTGGAACAGAACATCGCTGCCCT	300
Db	271	AATCATTTCTTCACCAACTGCTGCATGTCCCTGTGGGTGGAACAGAACATCGCTGCCCT	330
QY	301	CTCTGCCACGACGACGCTGGGTGGTCCAAAGAAATCGGCAATAGAGTGGTTGAAGGCTTC	360
Db	331	CTCTGCCACGACGACGCTGGGTGGTCCAAAGAAATCGGCAATAGAGTGGTTGAAGGCTTC	390
QY	361	TTACGCGAGTTGTCAGAGCCCTGGTGGATCTTGTAAATCCAGTCCCTCAACAAAGGCTTGA	420
Db	391	TTACGCGAGTTGTCAGAGCCCTGGTGGATCTTGTAAATCCAGTCCCTCAACAAAGGCTTGA	450
QY	421	ACACTACAGGGGAGTAATTTCTCAAAATAGAGCCGATGGATCTGTGTC--TTTGACTC	478
Db	451	ACACTACAGGGGAGTAATTTCTCAAAATAGAGCCGATGGATCTGTGTC--TTTGACTC	510
QY	479	ATCAAGAGCTTGGT--AGCATTTGTCTGTTTATCTCAGAAATCTCTGATTAAGA	536
Db	511	ATCAAGAGCTTGGT--AGCATTTGTCTGTTTATCTCAGAAATCTCTGATTAAGA	570
QY	537	AGATAAATTTTAA--GGTGGTCTTCTTACCTCTGTGGTGTGTCGCGACACAGCTT	595
Db	571	AGATAAATTTTAA--GGTGGTCTTCTTACCTCTGTGGTGTGTCGCGACACAGCTT	630
QY	596	AGAAGTGTATAAAAAGGAAGAGCTCAAAATGAATCACCTTATATTAATTAACCATTT	654
Db	631	AGAAGTGTATAAAAAGGAAGAGCTCAAAATGAATCACCTTATATTAATTAACCATTT	690
QY	655	CTATACAACAGCAGTGGAGACAGTTTC--GAGACTTTTTCGATGCTTATGGTGTATCAG	712
Db	691	CTATACAACAGCAGTGGAGACAGTTTC--GAGACTTTTTCGATGCTTATGGTGTATCAG	750
QY	713	TTA--AAAAAGATGTTACAGTAACAAATAAAGTGCATTTAAA	754
Db	751	TTAACAACAAGATGTTCCAGTAACAAATAAAGTGCATTTAAA	794
RESULT	4		
LOCUS	B1759082	918 bp	mRNA linear EST 25-SEP-2001
DEFINITION	603042879F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5183419 5',		
ACCESSION	B1759082		
VERSION	B1759082.1	GI:15750660	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (Bases 1 to 918)		
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/ .		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished. (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgabs-remail.nih.gov		
	Tissue Procurement: Life Technologies, Inc.		
	cDNA Library Preparation: Life Technologies, Inc.		
	cDNA library Arrayed by: The I.M.A.G.E. Consortium (LNL)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LNL at:		
	http://image.lnl.gov		

Plate: LHAM1457 row: m column: 20
High quality sequence stop: 732.

FEATURES

source

Location/Qualifiers

1..918

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_image="5183419"

/clone_id="NIH_MGC_116"

/lab_host="DH10B"

/note="Organ: pooled colon, kidney, stomach; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 3 colons, age 26 yo male, 49 yo female, 71 yo male colon; 46 yo male kidney, and pool of 2 stomachs, 62 yo male and 70 yo female. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 023. Note: this is a NIH_MGC Library."

BASE COUNT 269 a 196 c 247 g 206 t
ORIGIN

Query Match

Best Local Similarity 96.7%; Score 659.8; DB 10; Length 918;
Matches 728; Conservative 0; Mismatches 17; Indels 8; Gaps 5;

```

OY 1 ATGGCCGACGGGAGGAGGAGGAGAAACCTGGCCCTCTCTCAGTCCGGAGCTCA 60
DB 12 ATGGCCGACGGGAGGAGGAGGAGAAACCTGGCCCTCTCTCAGTCCGGAGCTCA 71
OY 61 GGCTCCAAGTCGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
DB 72 GGCTCCAAGTCGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 131
OY 121 TGGAGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180
DB 132 TGGAGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 191
OY 181 TGGCTTATGATGTCAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240
DB 192 TGGCTTATGATGTCAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 251
OY 241 AATCATTCCTTCCAACTCTGATGTCCTGCTGGGAGGAGGAGGAGGAGGAGGAG 300
DB 252 AATCATTCCTTCCAACTCTGATGTCCTGCTGGGAGGAGGAGGAGGAGGAGGAG 311
OY 301 CTCTGCCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
DB 312 CTCTGCCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 371
OY 361 TTAGCGAGTGTTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420
DB 372 TTAGCGAGTGTTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 431
OY 421 ACACTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 478
DB 432 ACACTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 491
OY 479 ATCAAAGCCTTGTT--AGCATTTGTCAGTTTATCTTCAGAAATTCCTGTGATTA 536
DB 492 ATCAAAGCCTTGTTAGCATTTTGTCAAGTTTATCTTCAGAAATTCCTGTGATTA 551
OY 537 AGATAATTTATTAAGTGTGCTCTCTACCTCTGCTGCTGCTGCTGCTGCTGCTG 596
DB 552 AGATAATTTATTAAGTGTGCTCTCTACCTCTGCTGCTGCTGCTGCTGCTGCTG 611
OY 597 GAAGTGTATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 655
DB 612 GAAGTGTATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 671
OY 656 TATACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 713

```

DB

672 TATACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 731

OY

714 TAA-AAAAGATGTTACAGTACCAATTAAGTG 745

DB

732 TAAACAAAAGATGTTACAGTACCAATTAAGTG 764

RESULT 5

Bi668735

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Bi668735 822 bp mRNA linear EST 12-SEP-2001
603293174F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5312226 5',
mRNA sequence.
Bi668735
Bi668735.1 GI:15582968
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 822)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs@mail.nih.gov
Tissue Procurement: Miklos Palokovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shihai
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAM1791 row: d column: 19
High quality sequence stop: 788.
Location/Qualifiers
1..822
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="5312226"
/clone_id="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescript (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTVNA-3',
size-selected for average insert size 2.3 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."

FEATURES

source

BASE COUNT 219 a 182 c 219 g 202 t
ORIGIN

Query Match 86.1%; Score 649; DB 10; Length 822;
Best Local Similarity 96.8%; Pred. No. 5.2e-169;
Matches 737; Conservative 0; Mismatches 15; Indels 9; Gaps 7;

```

OY 1 ATGGCCGACGTGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 60
DB 29 ATGGCCGACGTGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 88
OY 61 GGCTCCAAGTCGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
DB 89 GGCTCCAAGTCGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 148
OY 121 TGGAGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180
DB 149 TGGAGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 208

```

```

OY 181 TGTCTTAGATGTCAGCTGTAACAAACAAGAGAGCTGTGTGTGTGTGTGGGAGATGT 240
    |||||||
Db 209 TGTCTTAGATGTCAGCTGTAACAAACAAGAGAGCTGTGTGTGTGTGTGGGAGATGT 268
OY 241 AATCATTTCTTCACAACTGCTGCATGTCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300
    |||||||
Db 269 AATCATTTCTTCACAACTGCTGCATGTCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 328
OY 301 CTCTGTCACAGAGCTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 360
    |||||||
Db 329 CTCTGTCACAGAGCTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 388
OY 361 TTAGCCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420
    |||||||
Db 389 TTAGCCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 448
OY 421 ACACCTACAGGAGTGAATTTCTTCAATTAGAGCCGATGATGTGTGTGTGTGTGTGTGT 478
    |||||||
Db 449 ACACCTACAGGAGTGAATTTCTTCAATTAGAGCCGATGATGTGTGTGTGTGTGTGTGT 508
OY 479 ATCAAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 536
    |||||||
Db 509 ATCAAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 568
OY 537 AGATAATTTATTAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 596
    |||||||
Db 569 AGATAATTTATTAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 628
OY 597 GAAGTCTATTAATAA-AGCAAGAGCTCCAAATTAATGATACC-TTATTAATTAATTAAT 654
    |||||||
Db 629 GAAGTCTATTAATAA-AGCAAGAGCTCCAAATTAATGATACC-TTATTAATTAATTAAT 688
OY 655 CTATACAGAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 713
    |||||||
Db 689 CTATACAGAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 748
OY 714 TAAAAAGATGTTACAGTAACAATAAATGAAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 754
    |||||||
Db 749 T-ACCAAGATGTTACAGTAACAAT-AGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 787

```

```

RESULT 6
BI457840 818 bp mRNA linear EST 21-AUG-2001
LOCUS 603198212F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5277723.5
DEFINITION mRNA sequence.
ACCESSION BI457840.1 GI:15248496
VERSION EST.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM human.

```

```

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE NIH-MGC http://mhc.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue procurement: Miklos Palokovits, M.D., Ph.D.
cDNA library preparation: Michael J. Brownstein (NHGRI), Shitaki
Toshiyuki and Piero Carninci (RIKEN)
cDNA library arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L14M11701 row: 9 column: 04
High quality sequence stop: 780.
Location/Qualifiers
1..818
/organism="Homo sapiens"
/db_xref="taxon:9606"

```

```

/clone="IMAGE:5277723"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pBluescript (modified
pBluescript KS+); Site:1: BamHI; Site:2: SalI-XhoI (gtcgag
); Oligo-dr primed using primer 5'-TTTGTGTGTGTGTGTGTGTGT-3',
size-selected for average insert size 2.3 kb and
normalized to 10^5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC library."
BASE COUNT 225 a 174 c 214 g 205 t
ORIGIN

```

```

Query Match 85.9%; Score 647.4; DB 10; Length 818;
Best Local Similarity 95.6%; Pred. No. 1.4e-168;
Matches 733; Conservative 0; Mismatches 21; Indels 13; Gaps 6;

```

```

OY 1 ATGGCCGAGCTGGAACAGAGAGAAACCTGGCCCTGCTCTACCTCCGGAGACTCA 60
    |||||||
Db 17 ATGGCCGAGCTGGAACAGAGAGAAACCTGGCCCTGCTCTACCTCCGGAGACTCA 76
OY 61 GGCTCCAGTGGGAGGAGGAGCAAGATGTTCTCCCTCAAGAGTGAACGCGGAGCCATG 120
    |||||||
Db 77 GGCTCCAGTGGGAGGAGGAGCAAGATGTTCTCCCTCAAGAGTGAACGCGGAGCCATG 136
OY 121 TGGAGCTGGAGAGCTGAGTGCATGATGAGGAGGATGAGGATGAGTATGATGCC 180
    |||||||
Db 137 TGGAGCTGGAGAGCTGAGTGCATGATGAGGAGGATGAGGATGAGTATGATGCC 196
OY 181 TGTCTTAGATGTCAGCTGTAACAAACAAGAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 240
    |||||||
Db 197 TGTCTTAGATGTCAGCTGTAACAAACAAGAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 256
OY 241 AATCATTTCTTCACAACTGCTGCATGTCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300
    |||||||
Db 257 AATCATTTCTTCACAACTGCTGCATGTCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 316
OY 301 CTCTGTCACAGAGCTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 360
    |||||||
Db 317 CTCTGTCACAGAGCTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 376
OY 361 TTAGCCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420
    |||||||
Db 377 TTAGCCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 436
OY 421 ACACCTACAGGAGTGAATTTCTTCAATTAGAGCCGATGATGTGTGTGTGTGTGTGTGTGT 478
    |||||||
Db 437 ACACCTACAGGAGTGAATTTCTTCAATTAGAGCCGATGATGTGTGTGTGTGTGTGTGTGT 496
OY 479 ATCAAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 536
    |||||||
Db 497 ATCAAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 556
OY 537 AGATAATTTATTAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 596
    |||||||
Db 557 AGATAATTTATTAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 616
OY 597 GAAGTCTATTAATAA-AGCAAGAGCTCCAAATTAATGATACC-TTATTAATTAATTAAT 655
    |||||||
Db 617 GAAGTCTATTAATAA-AGCAAGAGCTCCAAATTAATGATACC-TTATTAATTAATTAAT 676
OY 656 TTATAC-AACAGCAGTGAAGCAGTTTC---GAGACTTTTTCGATGATGATGATGATGATG 710
    |||||||
Db 677 TTATACAAAGCAGTGAAGCAGTTTCGATGATGATGATGATGATGATGATGATGATGATG 736
OY 711 AGTT---AAAAAGATGTTACAGTAACAATAAATGAAGTGTGTGTGTGTGTGTGTGTGTGT 754
    |||||||
Db 737 CAGTTAAAAAGATGTTACAGTAACAATAAATGAAGTGTGTGTGTGTGTGTGTGTGTGTGT 783

```

```

RESULT 7
LOCUS B6708518
DEFINITION 602670411F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4793091 5'
ACCESSION B6708518
VERSION B6708518.1 GI:13985940
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 805)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabds-remail.nih.gov
Tissue Procurement: Miklos Palokovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraaki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10672 row: f column: 04
High quality sequence stop: 740.
FEATURES
source
1..805
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:4793091"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.3 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 214 a 183 c 214 g 194 t
ORIGIN
Query Match 85.5%; Score 644.4; DB 10; Length 805;
Best Local Similarity 94.6%; Pred. No. 9, 8e-168;
Matches 723; Conservative 0; Mismatches 31; Indels 10; Gaps 5;
QY 1 ATGGCGACGTGGAAGACGAGAGAACTGGCCCTGCTCAGCTCCGGAGGTCA 60
|
|
|
Db 26 ATGGCGACGTGGAAGACGAGAGAACTGGCCCTGCTCAGCTCCGGAGGTCA 85
|
|
|
QY 61 GGCTCAAGTGGGAGGCGACAAAGATGTTCTCCCTCAAGAGTGAACGGCGTGCATG 120
|
|
|
Db 86 GGCTCAAGTGGGAGGCGACAAAGATGTTCTCCCTCAAGAGTGAACGGCGTGCATG 145
|
|
|
QY 121 TGGAGCTGGAGCTGAGAGTCCGATAGTGGCCATCTGCAGGCTCCAGTGTATGCC 180
|
|
|
Db 146 TGGAGCTGGAGCTGAGAGTCCGATAGTGGCCATCTGCAGGCTCCAGTGTATGCC 205
|
|
|
QY 181 TGTCTAGATGTCAGCTGAAAAACAAGAGAGACGTGTTGGTGGGGAGATGT 240
|
|
|
Db 206 TGTCTAGATGTCAGCTGAAAAACAAGAGAGACGTGTTGGTGGGGAGATGT 265
|
|
|
QY 241 AATCATCTCTCCCAACAAGTCTGCATGCTCTGGTGGTGAACAGAAATCCCTCCCT 300
|
|
|
Db 266 AATCATCTCTCCCAACAAGTCTGCATGCTCTGGTGGTGAACAGAAATCCCTCCCT 325

```

```

QY 301 CTCTGCCAGCAGACTGGGTGGTCCAAAGATGGCAATGAGTGGTTAGAGGCTTC 360
|
|
|
Db 326 CTCTGCCAGCAGACTGGGTGGTCCAAAGATGGCAATGAGTGGTTAGAGGCTTC 385
|
|
|
QY 361 TTACGCGAGTTGTTCCAGAGCCCTGGTGATCTTAAATCCAGTCCCTACAAAGCTAGA 420
|
|
|
Db 386 TTACGCGAGTTGTTCCAGAGCCCTGGTGATCTTAAATCCAGTCCCTACAAAGCTAGA 445
|
|
|
QY 421 AACTACAGGGGATGAATTTCTCAATAGAGACCGATGATCTGTGTC--TTTGAAGTC 478
|
|
|
Db 446 AACTACAGGGGATGAATTTCTCAATAGAGACCGATGATCTGTGTC--TTTGAAGTC 505
|
|
|
QY 479 ATCAAGCCTTGGT--AGCATTTGTCAGTTTATCTTCGAATTCCTGTGATTAGA 536
|
|
|
Db 506 ATCAAGCCTTGGT--AGCATTTGTCAGTTTATCTTCGAATTCCTGTGATTAGA 565
|
|
|
QY 537 AGATTAATTTAATA--AGTGGTCTCTCTACCTCTGTGTGTGTGCGCACACACTT 595
|
|
|
Db 566 AGATTAATTTAATA--AGTGGTCTCTCTACCTCTGTGTGTGTGCGCACACACTT 625
|
|
|
QY 596 AGAAGTCTTTAAAGAAAGAGAGCTCAATTTGATACAC--TTATATTTACCATTT 654
|
|
|
Db 626 AGAAGTCTTTAAAGAAAGAGAGCTCAATTTGATACAC--TTATATTTACCATTT 685
|
|
|
QY 655 CTATACACAGCAGCAGTGGAGACAGTTTCGAG---ACTTTTGATGCTTATGTTGATC 710
|
|
|
Db 686 CTATACACAGCAGCAGTGGAGACAGTTTCGAG---ACTTTTGATGCTTATGTTGATC 745
|
|
|
QY 711 AGTTAAAGAAATGTTTACAGTAACAATAAATGTCAGTTTAA 754
|
|
|
Db 746 AGTTAAAGAAATGTTTACAGTAACAATAAATGTCAGTTTAA 789
|
|
|
RESULT 8
LOCUS B1601855
DEFINITION 603244729F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5287014 5'
ACCESSION B1601855
VERSION B1601855.1 GI:15494794
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 856)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabds-remail.nih.gov
Tissue Procurement: Miklos Palokovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraaki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11725 row: j column: 07
High quality sequence stop: 820.
FEATURES
source
1..856
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:5287014"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',

```


[illegible]

COMMENT

Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: DCTD/DRP
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov,
 Plate: L1CM/62 row: e column: 12
 High quality sequence stop: 767.
 Location/Qualifiers

FEATURES

source

```

1..855
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3929459"
/clone_lib="NIH_MGC_9"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/Note="Organ: ovary; Vector: pOMB7; Site: 1; XhoI; Site: 2;
ECORI; cDNA made by oligo-dT priming. Directionally
cloned into EORI/XhoI sites using the following 5'
adapter: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

BASE COUNT 250 a 161 c 238 g 205 t 1 others

ORIGIN

Query Match 82.8%; Score 624.6; DB 10; Length 855;
 Best Local Similarity 97.6%; Pred. No. 3.1e-162;
 Matches 698; Conservative 0; Mismatches 9; Indels 8; Gaps 6;

```

QY 30 CTGCGCCCTGCGCTCTCACTCCGGAGCTCAGCTCCCAAGTGGGAGCGGACAGATGTT 89
    |||
DB 1 CTGCGCCCTGCGCTCTCACTCCGGAGCTCAGCTCCCAAGTGGGAGCGGACAGATGTT 60
    |||
QY 90 CTCCTCAAGAAGTGAAGCGGCTGCGCATGTGGAGCTGGGACGTGGAGTGGATACGTG 149
    |||
DB 61 CTCCTCAAGAAGTGAAGCGGCTGCGCATGTGGAGCTGGGAGCTGGAGTGGATACGTG 120
    |||
QY 150 CGCATCTCAGAGGCTCCAGGTATGATGCTCTCTTAAATGTCAAGTGAACAAACA 209
    |||
DB 121 CGCATCTCAGAGGCTCCAGGTATGATGCTCTCTTAAATGTCAAGTGAACAAACA 180
    |||
QY 210 AGAGCACTGTGTGTGTGTGGGAGATGTAATCAATCTCTCCACACTGCTGATGTC 269
    |||
DB 181 AGAGCACTGTGTGTGTGTGGGAGATGTAATCAATCTCTCCACACTGCTGATGTC 240
    |||
QY 270 CCTGTGGTGAAGAAACAATCGCTGCTGCGCCAGAGGAGTGGTGGTCCAAAG 329
    |||
DB 241 CCTGTGGTGAAGAAACAATCGCTGCTGCGCCAGAGGAGTGGTGGTCCAAAG 300
    |||
QY 330 AATCGCAATGAGATGTTAGAAAGCTTCTTAAAGGAGTGTTCAGAGCCCTGTGGA 389
    |||
DB 301 AATCGCAATGAGATGTTAGAAAGCTTCTTAAAGGAGTGTTCAGAGCCCTGTGGA 360
    |||
QY 390 TCTTGAATCCAGTCCCTACAAAGGCTTGAACACTACAGGGAGCATTTCTCAATAG 449
    |||
DB 361 TCTTGAATCCAGTCCCTACAAAGGCTTGAACACTACAGGGAGCATTTCTCAATAG 420
    |||
QY 450 GAGCCGATGATCTGTGTGTC--TTTGGACTCATCAAGGCTTGTACCATTTGCACT 506
    |||
DB 421 GAGCCGATGATCTGTGTGTC--TTTGGACTCATCAAGGCTTGTACCATTTGCACT 480
    |||
QY 507 TTTATCTTCAAGAAATCTCTGTGATTAAGAAGTAAATTTAAAGTGGTCTTCTCTAC 566
    |||
DB 481 TTTATCTTCAAGAAATCTCTGTGATTAAGAAGTAAATTTAAAGTGGTCTTCTCTAC 540
    |||
QY 567 CTCTGTGTGTGTGTGTGGGAGCAGACAGCTTGAAGAGTCTTAAAAAAGAAAGACTCCA 626
    |||
DB 541 CTCTGTGTGTGTGTGTGGGAGCAGACAGCTTGAAGAGTCTTAAAAAAGAAAGACTCCA 600
    |||
```

```

QY 627 ATTGAATCACC--TTAATATTACCATTTCTATACAAC--GGCAGTGGAGAGTTTC-- 682
    |||
DB 601 ATTGAATCACCCTTTTAATTTACCCATTTCTATACAACGGGAGTGGAGAGTTTCAG 660
    |||
QY 683 GAGACTTTTTCAGATGCTTATGTTGCTTATGATGATGATGATGATGATGATGATGAT 737
    |||
DB 661 AAGACTTTTTCAGATGCTTATGTTGATGATGATGATGATGATGATGATGATGATGAT 714
    |||
```

RESULT 12

LOCUS

BG037017 893 bp mRNA linear EST 24-JAN-2001

DEFINITION

602287310P1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4374483 5',

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Human.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)

Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov

Tissue Procurement: Miklos Palakoyts, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shikaki
 Toshiyuki and Piero Carninci (RIKEN)

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: L1AM10038 row: d column: 04

High quality sequence stop: 741.

FEATURES

source

```

1..893
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4374483"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/Note="Organ: Brain; Vector: pBluescript (modified
pBluescript KS+); Site: 1; BamHI; Site: 2; SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.3 kb and
normalized to R0.5. This is a primary library enriched
for full-length clones and constructed using the
cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
```

BASE COUNT

ORIGIN

Query Match 82.7%; Score 623.6; DB 10; Length 893;
 Best Local Similarity 97.1%; Pred. No. 6e-162;
 Matches 678; Conservative 0; Mismatches 14; Indels 6; Gaps 4;

```

QY 1 ATGGCCGAGCTGGAAGAGCAGAGAAACCTGCGCCCTGCTCATCCTCCGGAGCTCA 60
    |||
DB 29 ATGGCCGAGCTGGAAGAGCAGAGAAACCTGCGCCCTGCTCATCCTCCGGAGCTCA 88
    |||
QY 61 GGCCTCAAGTGGGAGGAGCAGCAAGATGTTCTCCCTCAAGAAAGTGAAGCGGCTGCATG 120
    |||
DB 89 GGCCTCAAGTGGGAGGAGCAGCAAGATGTTCTCCCTCAAGAAAGTGAAGCGGCTGCATG 148
    |||
QY 121 TGGAGCTGGAGCTGAGTACGTGCGCATCTGACGAGGCTCAGGTGATGATGCC 180
    |||
```

```

Db 149 TGGAGCTGGAGCTGAGTGCATACGTGCGCCATCTGACGGGTCCAGGTGATGATGCC 208
Oy 181 TGTCTTAGATGTCAAGCTGAAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
Db 209 TGTCTTAGATGTCAAGCTGAAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 268
Oy 241 ATTCATTCCTTCCAACTGCTGCATGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
Db 269 ATTCATTCCTTCCAACTGCTGCATGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 328
Oy 301 CTCTCCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
Db 329 CTCTCCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 388
Oy 361 TTAGCGCAGTGTTCAGAGCCCTGCTGATCTGTATTCAGAGCCCTACAAAGAGCTAGA 420
Db 389 TTAGCGCAGTGTTCAGAGCCCTGCTGATCTGTATTCAGAGCCCTACAAAGAGCTAGA 448
Oy 421 ACACCTACAGGGAGATGATTCCTCAAAATAGAGCCGATGATCTGTGTC--TTTGGACTC 478
Db 449 ACACCTACAGGGAGATGATTCCTCAAAATAGAGCCGATGATCTGTGTC--TTTGGACTC 508
Oy 479 ATCAAGGCTTGTGTT--AGCATTTGTCTGTTTATCTTCAGAAATTCCTGTGATTAAGA 536
Db 509 ATCAAGGCTTGTGTTTACATTTTGTCTGATTTTATCTTCAGAAATTCCTGTGATTAAGA 568
Oy 537 AGATTAATTTATTA--AGGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 595
Db 569 AGATTAATTTATTA--AGGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 628
Oy 596 AGAAGTGTATTAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 654
Db 629 AGAAGTGTATTAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 688
Oy 655 CTATCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 692
Db 689 CTATCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 726

```

```

RESULT 13
Bi858307 797 bp mRNA linear EST 10-OCN-2001
LOCUS 603384080F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:5393131 5'
DEFINITION mRNA sequence.
ACCESSION Bi858307
VERSION Bi858307.1 GI:15999054
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
            Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 797)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC).
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
            Email: cgapbs@mail.nih.gov
            Tissue Procurement: DCTD/DTF
            cDNA Library Preparation: Life Technologies, Inc.
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLNL2001 row: 0 column: 20
            High quality sequence stop: 718.
            Location/Qualifiers
                1..797
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:5393131"
                /clone_1ib="NIH_MGC_87"
                /tissue_type="mammary adenocarcinoma, cell line"

```

```

FEATURES
    source

```

```

/lab_host="dh10b (phage-resistant)"
/note="Organ: breast; Vector: pCMV-SPORT6; Site_1: NotI;
      Site_2: SalI; Cloned unidirectionally; 0.190-0.191 kb.
      Average insert size 1.383 kb. Library enriched for
      full-length clones and constructed by Life Technologies.
      Note: this is a NIH-MGC Library."
BASE COUNT 229 a 166 c 204 g 198 t
ORIGIN

```

```

Query Match 81.7%; Score 616.2; DB 10; Length 797;
Best Local Similarity 95.0%; Pred. No. 6,5e-160;
Matches 725; Conservative 0; Mismatches 23; Indels 15; Gaps 8;

```

```

Oy 4 GCCACGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 63
Db 1 GCCACGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
Oy 64 TCCAGTTCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 123
Db 61 TCCAGTTCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
Oy 124 AGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 183
Db 121 AGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
Oy 184 CTATGATGTCAAGCTGAAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 243
Db 181 CTATGATGTCAAGCTGAAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
Oy 244 CATTCCTTCCCAACATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 303
Db 241 CATTCCTTCCCAACATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
Oy 304 TGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 363
Db 301 TGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
Oy 364 GCGCAGTGTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 423
Db 361 GCGCAGTGTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Oy 424 CTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 481
Db 421 CTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
Oy 482 AAGCCTTGTGTT--AGCATTTGTCAATTTATCTTCAAGAAATTCCTGCGATTAAAGAA 539
Db 481 AAGCCTTGTGTTTACGATTTTGTCAATTTATCTTCAAGAAATTCCTGCGATTAAAGAA 540
Oy 540 TAATTTATTT--AAGAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 598
Db 541 TAATTTATTTCAAGAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 600
Oy 599 AGTGTATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 656
Db 601 AGTGTATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
Oy 657 ATCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 714
Db 661 AT---ACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 717
Oy 715 AAA---AAGAAATGTTACAGTAACAAATTAAGTCAAGTTTAA 754
Db 718 AAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 760

```

```

RESULT 14
Bi753323 856 bp mRNA linear EST 15-MAY-2001
LOCUS 602731740F1 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:4875489 5'
DEFINITION mRNA sequence.
ACCESSION Bi753323

```

```

VERSION      BG753323.1  GI:14063976
KEYWORDS     EST.
SOURCE       human.
ORGANISM     Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 856)
              NIH-MGC http://mgs.nci.nih.gov/.
              National Institutes of Health, Mammalian Gene Collection (MGC).
              Unpublished (1999)
AUTHORS      Contact: Robert Strausberg, Ph.D.
              Email: cga@bbs-rmail.nih.gov
              Tissue Procurement: ATCC
              CDNA Library Preparation: Ling Hong/Rubin Laboratory
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: L1CM1756 row: 0 column: 10
              High quality sequence stop: 767.
FEATURES     location/Qualifiers
              1..856
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /clone="IMAGE:4875489"
              /clone_lib="NIH_MGC_43"
              /tissue_type="normal pigmented retinal epithelium".
              /lab_host="DH10B (phage-resistant)".
              /note="Organ: eye; Vector: pOTR8; Site_1: XhoI; Site_2:
              EcoRI; CDNA made by oligo-dT priming. Directionally
              cloned into EcoRI/XhoI sites using the following 5'
              adaptor: GGCACGAG(G). Library constructed by Ling Hong
              in the laboratory of Gerald M. Rubin (University of
              California, Berkeley) using ZAP-CDNA Synthesis Kit
              (Stratagene) and Superscript II RT (Life Technologies).
              Note: this is a NIH_MGC Library. |"
BASE COUNT   231 a      214 c      211 g      200 t
ORIGIN
Query Match      80.8%; Score 609; DB 10; Length 856;
Best Local Similarity 94.1%; Pred. No. 6,6e-158;
Matches 712; Conservative 0; Mismatches 30; Indels 15; Gaps 7;
QY 1 ATGGCCGAGCTGGAAGAGAGAGAACTGGCCCTCTCTCACTCGGGAGCTCA 60
    |||||||
DB 18 ATGGCCGAGCTGGAAGAGAGAGAACTGGCCCTCTCTCACTCGGGAGCTCA 77
QY 61 GGCTCCAGTCGGGAGGCGAAGATGTTCTCCCTCAAGAAAGTGAACGGGGGCGCATG 120
    |||||||
DB 78 GGCTCCAGTCGGGAGGCGAAGATGTTCTCCCTCAAGAAAGTGAACGGGGGCGCATG 137
QY 121 TGGAGCTGGAGCTGAGTGCATGAGTGGCCATCTGACGAGGTCCAGTGAATGATGCC 180
    |||||||
DB 138 TGGAGCTGGAGCTGAGTGCATGAGTGGCCATCTGACGAGGTCCAGTGAATGATGCC 197
QY 181 TGTCTTAGATGTCAGAGTGAACAAACAAGAGAGACTGTGTGTGCTGGGAGATGT 240
    |||||||
DB 198 TGTCTTAGATGTCAGAGTGAACAAACAAGAGAGACTGTGTGTGCTGGGAGATGT 257
QY 241 AATCATTTCTTCCACAACCTGCTCATGTCCTGCTGGTGAACAACAATCGCTGCCCT 300
    |||||||
DB 258 AATCATTTCTTCCACAACCTGCTCATGTCCTGCTGGTGAACAACAATCGCTGCCCT 317
QY 301 CTCCTGCAGCAGACTGTGGTGTCTCAAGAATCGCCAAATGAGATGTTAGAAGCTTC 360
    |||||||
DB 318 CTCCTGCAGCAGACTGTGGTGTCTCAAGAATCGCCAAATGAGATGTTAGAAGCTTC 377
QY 361 TTAGCGCAGTTGTCAGAGCCCTGTGGATCTTGTAACTCAGTGCCTTACAAAGCTAGA 420
    |||||||
DB 378 TTAGCGCAGTTGTCAGAGCCCTGTGGATCTTGTAACTCAGTGCCTTACAAAGCTAGA 437

```

```

QY 421 ACACACAGGGGATGATCTTCCAAATAGAGCCGATGATCTGTGTC- -TTTGACATC 478
    |||||||
DB 438 ACACACAGGGGATGATCTTCCAAATAGAGCCGATGATCTGTGTC- -TTTGACATC 497
QY 479 ATCAAGCCCTTGTT- -AGCATTTGTCAGTTTATCTTCAGAAATTCCTGTGATTAAGA 536
    |||||||
DB 498 ATCAAGCCCTTGTTAGCATTTTGTGACATTTTATCTTCAGAAATTCCTGTGATTAAGA 557
QY 537 AGATTAATTTATTA- -AGGTGTCCTTCCTACCTGCTGGGTGTGTGCGCACACAGCTT 595
    |||||||
DB 558 AGATTAATTTATTA- -AGGTGTCCTTCCTACCTGCTGGGTGTGTGCGCACACAGCTT 617
QY 596 AGAAGTG-CTATTAAGAAAGAGAGAGCTCCAAATGATCACC- -TTATATTTACCAT 653
    |||||||
DB 618 AGAAGTGCTATTAAGAAAGAGAGAGCTCCAAATGATCACC- -TTATATTTACCAT 677
QY 654 TCTATACACAGCAGCTGGAAGCAGTTTGAGACTTTTCGATGC- -TTATGTTG 707
    |||||||
DB 678 TCTATACACAGCAGCTGGAAGCAGTTTGAGACTTTTCGATGC- -TTATGTTG 737
QY 708 ATCAGTTAA- -AAAGATGTTACAGTAAACAAATTA 742
    |||||||
DB 738 ATCAGTTAA- -AAAGATGTTACAGTAAACAAATTA 774
RESULT 15
BI601470 767 bp mRNA linear EST 07-SEP-2001
LOCUS 603249053F1 NIH_MGC_96 Homo sapiens CDNA clone IMAGE:5300845 5',
DEFINITION mRNA sequence.
ACCESSION BI601470
VERSION BI601470.1 GI:15494409
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
              1 (bases 1 to 767)
              NIH-MGC http://mgs.nci.nih.gov/.
              Unpublished (1999)
              National Institutes of Health, Mammalian Gene Collection (MGC)
              Contact: Robert Strausberg, Ph.D.
              Email: cga@bbs-rmail.nih.gov
              Tissue Procurement: Miklos Palovits, M.D., Ph.D.
              CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
              Toshiyuki and Piero Carninci (RIKEN)
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: L1AM11761 row: 1 column: 14
              High quality sequence stop: 701.
FEATURES     location/Qualifiers
              1..767
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /clone="IMAGE:5300845"
              /clone_lib="NIH_MGC_96"
              /tissue_type="hypothalamus"
              /lab_host="DH10B"
              /note="Organ: brain; Vector: pBluescript (modified
              pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
              ); Oligo-dT primed using primer 5'-TTTATTTTATTTTATTTATTT-3',
              size-selected for average insert size 2.3 kb and
              normalized to 50x. This is a primary library enriched
              for full-length clones and constructed using the
              Cap-trapper method (Carninci, in preparation) library
              constructed by M. Brownstein (NHGRI/NHGRI, National
              Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT   192 a      173 c      214 g      188 t
ORIGIN

```

Query Match	80.7%;	Score 608.8;	DB 10;	Length 767;
Best Local Similarity	94.2%;	Pred. No. 7.2e-158;		
Matches 688;	Conservative	0;	Mismatches 32;	Indels 10;
				Gaps 5;

[illegible]

Search completed: July 25, 2002, 18:05:48
Job time: 14227 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 25, 2002, 14:08:41 ; Search time 2540.26 Seconds
(without alignments)
6057.068 Million cell updates/sec

Title: US-09-509-779-1

Perfect score: 1140

Sequence: 1 GTTCTGGCGCCGCCATG.....ACTAATTCATCAATTAATG 1140

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estbta:*
2: em_estbta:*
3: em_estin:*
4: em_estin:*
5: em_estin:*
6: em_estov:*
7: em_estov:*
8: em_estro:*
9: em_estro:*
10: em_estro:*
11: em_estro:*
12: em_estro:*
13: em_gss_hum:*
14: em_gss_hum:*
15: em_gss_hum:*
16: em_gss_vfl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1069.4	93.8	1132	11 AK003963	AK003963 Mus muscu
2	1041	91.3	1114	11 AK011328	AK011328 Mus muscu
3	1032.4	90.6	1084	11 AK003248	AK003248 Mus muscu
4	1004.4	88.1	1078	11 AK002628	AK002628 Mus muscu
5	765	67.1	820	10 BI732003	BI732003 603355869
6	753.4	66.1	1079	10 BI410936	BI410936 602962128
7	700.8	61.5	722	11 AK007588	AK007588 Mus muscu
8	686.6	60.2	731	10 BI647378	BI647378 603279634
9	673.4	59.1	978	10 BI156184	BI156184 602903365
10	665.4	58.4	916	10 BG917948	BG917948 602820877
11	651.4	57.1	754	10 BI147544	BI147544 602913971
12	651	57.1	704	10 BF018804	BF018804 ux83b03.x
13	649.2	56.9	946	10 BF383735	BF383735 602044644
14	646.2	56.7	737	10 BI556918	BI556918 603339970
15	624.4	54.8	822	10 BF302267	BF302267 602031233
16	612.2	53.7	755	10 BF579825	BF579825 602095851
17	597.2	52.4	620	10 BE628111	BE628111 uu29a11.x

c	18	596	52.3	651	9	AM555800	AM555800 L0259F11-
	19	575	50.4	575	10	BF023423	BF023423 ux08a05.y
	20	568.8	49.9	575	10	BI153818	BI153818 602870803
	21	567	49.7	568	10	BF720947	BF720947 mab59f05.
	22	556.8	48.8	577	9	AA230335	AA230335 mw04c09.r
	23	554.8	48.7	574	10	BG087875	BG087875 H3146A01-
	24	547.2	48.0	576	9	AA221462	AA221462 my23c09.r
	25	542.2	47.6	760	10	BF163003	BF163003 601767409
	26	542.2	47.5	558	9	AM542025	AM542025 C0153G11-
	27	537.4	47.1	852	9	BF159848	BF159848 601767509
	28	536.4	47.1	627	11	AK006703	AK006703 Mus muscu
	29	530.6	46.5	847	10	BF140544	BF140544 601787577
	30	525.6	46.1	881	10	BE915897	BE915897 601669166
	31	522.4	45.8	589	10	BC522883	BC522883 ux62d04.x
	32	522	45.8	522	10	BE372989	BE372989 601224463
	33	519	45.5	620	9	AM989997	AM989997 uf31b08.y
	34	518.6	45.5	687	10	C88520	C88520 C88520 Mous
	35	513	45.0	562	10	BE824793	BE824793 C0504H10-
	36	512.4	44.9	563	10	BE143878	BE143878 ut66a03.x
	37	512	44.9	513	10	BE850994	BE850994 uw91b02.y
	38	505.4	44.3	538	9	AM491099	AM491099 UI-M-BH3-
	39	504.8	44.3	922	10	BF161048	BF161048 601767681
	40	502	44.0	513	9	AA032932	AA032932 mi22d08.r
	41	501.8	44.0	542	9	AA475269	AA475269 vH20b08.r
	42	501	43.9	501	10	BE628851	BE628851 uu29a11.y
	43	494	43.3	494	10	BC277865	BC277865 ux62d04.y
	44	493	43.2	555	10	BF730110	BF730110 mab59f05.
	45	491	43.1	507	10	BM022893	BM022893 id63d03.y

ALIGNMENTS

RESULT 1	AK003963	1132 bp	mus musculus 18 days embryo whole body cDNA, RIKEN full-length enriched library, clone:1110029M05:ring finger protein 7, full insert sequence.
LOCUS	AK003963		
DEFINITION	AK003963.1 GI:12834939		
ACCESSION	AK003963		
VERSION	HTC; CAP trapper.		
KEYWORDS	Mus musculus (strain:C57BL/6J) 18 days embryo cDNA to mRNA, clone.lib:RIKEN full-length enriched mouse cDNA library		
SOURCE	clone:1110029M05.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	Carninci, P. and Hayashizaki, Y.		
TITLE	High-efficiency full-length cDNA cloning		
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)		
MEDLINE	992979253		
PUBMED	10349636		
REFERENCE	2 (sites)		
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes		
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)		
MEDLINE	20499374		
PUBMED	11042159		
REFERENCE	3 (sites)		
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashtanov, I., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, Y., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.		
TITLE	RIKEN integrated sequence analysis (RISA) system--384-Format		
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)		

[illegible]

	Encyclopedic Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGACGATCCCAAGACCTCTTTTCTTTTTTAAATGYN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 7.5 and subtraction to Rot = 37.5. Second strand cDNA was prepared with the primer adpotele of sequence [5' GAGAGAGACATTCCTCAGTTAATTAAATTAAATCCCCCCCCCCC 3'], cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI. Host: SOLR.					
FEATURES						
source	Location/Qualifiers					
	1..1114					
	/organism="Mus musculus"					
	/strain="C57BL/6J"					
	/db_xref="MGD:MGI:1893659"					
	/db_xref="taxon:10090"					
	/clone="2610005N2"					
	/clone_lib="RIKEN full-length enriched mouse cDNA library"					
	/dev_stage="10 days embryo"					
gene	1..1114					
	/gene="Rnf7"					
misc_feature	1..1114					
	/gene="Rnf7"					
	/note="data source:MGD, source key:MGI:1337096, evidence:ISS					
	ring finger protein 7"					
BASE COUNT	289 a 220 c 287 g 318 t					
ORIGIN						
Query Match	91.3% Score 1041; DB 11; Length 1114;					
Best Local Similarity	99.3%; Pred. No. 3.7e-227;					
Matches 1109; Conservative	0; Mismatches 0; Indels 8; Gaps 6;					
OY	4	CTGGCGCCGCCGCATGGCGCGCATGTGGAGAGGAGGAGAACCCCTGCGTCTTCTTGCGA	63			
Db	2	CTGGCGCCGCCGCATGGCGCGCATGTGGAGAGGAGGAGAACCCCTGCGTCTTCTTGCGA	60			
OY	64	CTCCGGGAGCGCAGGCTCCAAGTCGGAGGCGCACAAATGTTCTCTCAAGAAGTGGA	123			
Db	61	CTCCGGGAGCGC - GGCTCCAAGTCGGAGGCGCACAAAGTGTCTCTCAAGAAGTGGA	119			
OY	124	C CGCGT AACCAT GTG GACT GGG AGC GTT GA TGA TG C CAT FAC CTG TGC CAT CTC TGA AGG CT CA	183			
Db	120	C GCGGT ACCCAT GTG TGA ACT TGG AAG GTT GA TGA TG C CAT FAC CTG TGC CAT CTC TGA AGG CT CA	179			
OY	184	G GTG ATG ATG C TC TC CC TT GC AT GTG CA AG CT AA A A CA AG CA AG AG AC TG GTT GT GT	243			
Db	180	G GTG ATG ATG C TC TC CC TT GC AT GTG CA AG CT AA A A CA AG CA AG AG AC TG GTT GT GT	239			
OY	244	CTGGGAGAGTGTAAACAATTCCTTCCACAACTGTGCATGTCCCTGTGGGTGAACAGAA	303			
Db	240	CTGGGAGAGTGTAAACAATTCCTTCCACAACTGTGCATGTCCCTGTGGGTGAACAGAA	299			
OY	304	CAATGCGTGGCCCTGTGTGCGACGACAGACATGGGTATCTCAAAAGAAATCGCAATGTAGAGT	363			
Db	300	CAATGCGTGGCCCTGTGTGCGACGACAGACATGGGTATCTCAAAAGAAATCGCAATGTAGAGT	359			
OY	364	G GCCC AGC GCG TCC TGT GTG TGG TTT CCT CTA ACC CTG GACA AAA GCA CTT AAG CACT GAG GGA T	423			
Db	360	G GCCC AGC GCG TCC TGT GTG TGG TTT CCT CTA ACC CTG GACA AAA GCA CTT AAG CACT GAG GGA T	419			
OY	424	T CAT C CT TGA G A G A G A G A G A T G T G T G G C C T T T T G A G A C T C A C C A A A G C T T G C T T A T	483			
Db	420	T CAT C CT TGA G A G A G A G A G A T G T G T G G C C C T T T T G A G A C T C A C C A A A G C C T T G C T T A T	479			
OY	484	T A A T T T G C T G T T T G T T T T G G A A A T C T C A C A A T T A A G A T A T T T G T T A A A A T G G C	543			
Db	480	T A A T T T G C T G T T T G T T T T G G A A A T C T C A C A A T T A A G A T A T T T G T T A A A A T G G C	539			

/translation="MADVEDGEPCVLSHSGSAGSKSGDKMFLKWNVAAMSMID
 VEDPCICRVOVMACLRQCAENKQEDCVVWVWGEQNHSHFHCMSIMWKQNNRCPIC
 QODWVYORIGK"
 polyA_signal 1062..1067 /note="putative"
 polyA_site 1084 /note="putative"
 BASE COUNT 277 a 218 c 285 g 304 t
 ORIGIN

Query Match 90.6%; Score 1032.4; DB 11; Length 1084;
 Best Local Similarity 99.4%; Pred. No. 3.4e-225;
 Matches 1079; Conservative 0; Mismatches 1; Indels 6; Gaps 4;

QY 4 CTCGCGCGCGCGCATGGCCGACGTGGAGACGGCGAGAACCTCGCTCTTCTTCGCA 63
 |||
 DB 2 CTCGCGCGCGCGCATGGCCGACGTGGAGACGGCGAGAACCTCGCTCTTCTTCGCA 61
 |||
 QY 64 CTCGCGGAGCGCAGGCTCCAAAGTCGGAGGAGCAAGATGTCTCTCAAGAAGTGAA 123
 |||
 DB 62 CTCGCGGAGCGCAGGCTCCAAAGTCGGAGGAGCAAGATGTCTCTCTCAAGAAGTGAA 121
 |||
 QY 124 CGCGGTAAGCCATGTGAGCTGGAGCTGGAGTTCGATACCTGTGCCATCTGCAAGGTCCA 183
 |||
 DB 122 CGCGGTAAGCCATGTGAGCTGGAGCTGGAGTTCGATACCTGTGCCATCTGCAAGGTCCA 181
 |||
 QY 184 GGTGATGATGATGCTGCTGCTGATGTCAAGTGAAGAACAGAGAGAGCTGTGTGTGT 243
 |||
 DB 182 GGTGATGATGATGCTGCTGCTGATGTCAAGTGAAGAACAGAGAGAGCTGTGTGTGT 241
 |||
 QY 244 CTCGGGAGAGTGAACCATCTCTCCCAACTGCTGCATGCTCCGTGGGTGAAGACGAA 303
 |||
 DB 242 CTCGGGAGAGTGAACCATCTCTCCCAACTGCTGCATGCTCCGTGGGTGAAGACGAA 301
 |||
 QY 304 CAATCGCTGCGCTCTGTGCCAGCAGAGACTGGTATGCCAAGAATCGCAATGAGAGGT 363
 |||
 DB 302 CAATCGCTGCGCTCTGTGCCAGCAGAGACTGGTATGCCAAGAATCGCAATGAGAGGT 361
 |||
 QY 364 GGGCCAGAGCGCTCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 423
 |||
 DB 362 GGGCCAGAGCGCTCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 421
 |||
 QY 424 TCATCCTTGAGAGAGAGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 483
 |||
 DB 422 TCATCCTTGAGAGAGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 481
 |||
 QY 484 TAAATGCTGT 543
 |||
 DB 482 TAAATGCTGT 541
 |||
 QY 544 CTTTCCCTACCTCGT 603
 |||
 DB 542 CTTTCCCTACCTCGT 601
 |||
 QY 604 AATGATCTTGT 660
 |||
 DB 602 AATGATCTTGT 661
 |||
 QY 661 GTTGT 720
 |||
 DB 662 GTTGT 721
 |||
 QY 721 ATGATGTGAAGAACCGCCTCTCTTAATCTCTTTTGTGTGTGTGTGTGTGTGTGTGTGT 780
 |||
 DB 722 ATGATGTGAAGAACCGCCTCTCTTAATCTCTTTTGTGTGTGTGTGTGTGTGTGTGTGT 781
 |||
 QY 781 CACCTGCTGTCTTCAATTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 840
 |||
 DB 782 CACCTGCTGTCTTCAATTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 841
 |||
 QY 841 TAACTGTGGGTAACTGTATATATGCGGTACTGTGCGGTAAACGCTTTGTCTCTGTAC 900
 |||

DB 842 TAACTGTGGGTAACTGTATATATGCGGTACTGTGCGGTAAACGCTTTGTCTCTGTAC 901
 |||
 QY 901 TTCTCCATCTTTTACTTTGGCCAGAGACCTGTGATTTGTCAACCATTTAGTCTTAAGAAC 960
 |||
 DB 902 TTCTCCATCTTTTACTTTGGCCAGAGACCTGTGATTTGTCAACCATTTAGTCTTAAGAAC 961
 |||
 QY 961 TGTTCCTGTTTTTTTGGCGAAGGTTGTATTTAGTTTACTCAAAAATATTAGTAGGAA 1020
 |||
 DB 962 TGTTCCTGTTTTTTTGGCGAAGGTTGTATTTAGTTTACTCAAAAATATTAGTAGGAA 1019
 |||
 QY 1021 ATGCTTACTAGTATTAACACTGAAGTTCATATGCAATGTTTTTAATAAATATTGTGCTT 1080
 |||
 DB 1020 AT-GCTTACTAGTATTAACACTGAAGTTCATATGCAATGTTTTTAATAAATATTGTGCTT 1078
 |||
 QY 1081 TGAGCTT 1086
 |||
 DB 1079 TGAGCT 1084

RESULT 4

AK002628

LOCUS 1078 bp mRNA linear HTC 19-JAN-2002
 DEFINITION Mus musculus adult male kidney cDNA, RIKEN full-length enriched

library, clone:0610013021:ring finger protein 7, full insert
 sequence.

ACCESSION AK002628.1 GI:12832752
 VERSION AK002628

KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (strain:C57BL/6J) adult male kidney cDNA to mRNA,
 clone:lib:RIKEN full-length enriched mouse cDNA library
 clone:0610013021.

ORGANISM

Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (sites)
 Carninci,P. and Hayashizaki,Y.
 High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)
 99279253

PUBMED 10349636
 REFERENCE

AUTHORS

2 (sites)
 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
 Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)
 20499374

JOURNAL

MEDLINE 11042159
 PUBMED 11042159
 REFERENCE

AUTHORS

3 (sites)
 Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
 Konno,H., Akiyama,J., Nishi,K., Kitsuina,T., Tashiro,H., Itoh,M.,
 Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
 Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Washiwagi,K.,
 Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watanishi,M.,
 Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J.,
 Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
 RIKEN integrated sequence analysis (RISA) system-384-format
 sequencing pipeline with 384 multicapillary sequencer

Genome Res. 10 (11), 1757-1771 (2000)
 20530913

JOURNAL MEDLINE 11076861
 PUBMED 11076861
 REFERENCE

AUTHORS

4 (sites)
 The Riken genome Exploration Research Group Phase II Team and the
 FANTOM Consortium.
 Functional annotation of a full-length mouse cDNA collection

Nature 409, 685-690 (2001)
 5 (bases 1 to 1078)

JOURNAL MEDLINE 11076861
 PUBMED 11076861
 REFERENCE

AUTHORS

5 (bases 1 to 1078)
 Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,
 Arakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Bult,C.,
 Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T.,
 Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hirooka,T., Hoti,F.,
 Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T.,
 Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S.,

REFERENCE 1 (bases 1 to 820)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: The Cepko Laboratory
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LHAM1923 row: 1 column: 17
 High quality sequence stop: 740.

FEATURES
 Source
 1. 820
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="IMAGE:5163056"
 /clone_lib="NIH_MGC_94"
 /tissue_type="retina"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: eye; Vector: pCMV-Sport6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 3.3 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH-MGC library."

BASE COUNT 219 a 157 c 204 g 240 t
 ORIGIN

Query Match 67.1%; Score 765; DB 10; Length 820;
 Best Local Similarity 98.2%; Pred. No. 3.4e-164;
 Matches 806; Conservative 0; Mismatches 10; Indels 5; Gaps 3;

224 CAAGAGACTGTGTGTGTGGGGAGAGTGAACCATTCCTCCACAACCTGCTGATG 283
 1 CAAGAGACTGTGTGTGTGGGGAGAGTGAACCATTCCTCCACAACCTGCTGATG 60
 284 TCCTGTGGGTGAACGAAACATCGTGCCTCTGTGCCAGACAGGCTGGTAGTCAA 343
 1 TCCTGTGGGTGAACGAAACATCGTGCCTCTGTGCCAGACAGGCTGGTAGTCAA 120
 61 TCCTGTGGGTGAACGAAACATCGTGCCTCTGTGCCAGACAGGCTGGTAGTCAA 120
 344 AGAATCGGCAATGAGAGGTGGCCAGAGCGCTGCTGTGTGTGTGTGTGTGTGTGT 403
 121 AGAATCGGCAATGAGAGGTGGCCAGAGCGCTGCTGTGTGTGTGTGTGTGTGTGT 180
 404 GACTAAACACTGAGGGGATTCATCTTGAAGAGAGAGAGAGTCTGTGCCCTTTGAGAC 463
 181 GACTAAACACTGAGGGGATTCATCTTGAAGAGAGAGAGAGTCTGTGCCCTTTGAGAC 240
 464 TCACCAAGGCTGCTTATTAATTTGTCTGTAGTTTGGGAATTCCTCTCAATTTAA 523
 241 TCACCAAGGCTGCTTATTAATTTGTCTGTAGTTTGGGAATTCCTCTCAATTTAA 300
 524 GATAATTTGTTAAATATGCTTCCACTCGGTGTGTGTGTGTGTGTGTGTGTGTGTGT 583
 301 GATAATTTGTTAAATATGCTTCCACTCGGTGTGTGTGTGTGTGTGTGTGTGTGTGT 360
 584 GAAGAGGAGAACACAGAAATGATCTTGTATCTGTACCCAGACT---GGAACAT 640
 361 GAAGAGGAGAACACAGAAATGATCTTGTATCTGTACCCAGACTGTGGGAACAT 420
 641 TGTGTTACAGAAACATTTGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 700
 421 TGTGTTACAGAAACATTTGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 480
 701 ATGTTACAGTAACAAATTAATGATTTGAAGAACCGACCTCTCTTAATCTTTTGTGT 760
 481 ATGTTACAGTAACAAATTAATGATTTGAAGAACCGACCTCTCTTAATCTTTTGTGT 540
 761 GGGAGAGAGGCAAGGAGGCCACCTGTGTCTTTCATTTGCTGTGAATGAGATTTTAAC 820

|||||
 Db 541 GGGAGAGGCAAGCGAGGCCACCCCTGCTTCATTTGCTGTGAATGAGATTTTAAC 600
 QY 821 CTCGACTCACTGAAGAGCGCTTAAGCTGATATGCGGTAACTGCTGGGT 880
 Db 601 CTCGACTCACTGAAGAGCGCTTAAGCTGATATGCGGTAACTGCTGGGT 660
 QY 881 AAGCGC-TTGTGCTCTGACTCTCCATCTTTGACTTGGCCAGAGCTGATTTTC 939
 Db 661 AAGCGCTTTGCTCTGACTCTCCATCTTTGACTTGGCCAGAGCTGATTTTC 720
 QY 940 AACCACTTAAGTAAGAACTGTTTCTGTTTGGCCAGAGCTGATTTTA 999
 Db 721 AACCACTTAAGTAAGAACTGTTTCTGTTTGGCCAGAGCTGATTTTA 779
 QY 1000 GTCAAAATTTAGTAGAAATGCGTTACTATTAAC 1040
 Db 780 GTCAAAATTTAGTAGAAATGCGTTACTATTAAC 820

RESULT 6
 BI410936 1079 bp mRNA linear EST 14-ANG-2001
 LOCUS 602962128F1 NCI_CGAP_Lu33 Mus musculus cDNA clone IMAGE:5117623 5',
 DEFINITION mRNA sequence.
 ACCESSION BI410936
 VERSION BI410936.1 GI:15171859
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 1079)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
 Bonaldo, Ph.D.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LHAM11286 row: h column: 08
 High quality sequence stop: 909.
 Location/Qualifiers
 1. 1079
 /organism="Mus musculus"
 /strain="C28H II"
 /db_xref="taxon:10090"
 /clone="IMAGE:5117623"
 /clone_lib="NCI_CGAP_Lu33"
 /tissue_type="pooled lung tumors"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: lung; Vector: pRT73D-Pac (Pharmacia) with a
 modified polylinker; Site_1: NotI; Site_2: EcoRI; 1st
 strand cDNA was prepared from mRNA obtained from pooled
 lung tumors with a Not I - oligo(dT) primer [5',
 TGTTCACATCTGACGTGAGCGGCGGCTCTGTTTGTGTGTGTGTGTGTGTGTGTGTGT
 3'].
 Double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not
 I and Eco RI sites of the modified pRT73 vector. Library
 went through one round of normalization, and was
 constructed by Bento Soares and M. Fatima Bonaldo."

FEATURES
 Source
 1. 1079
 /organism="Mus musculus"
 /strain="C28H II"
 /db_xref="taxon:10090"
 /clone="IMAGE:5117623"
 /clone_lib="NCI_CGAP_Lu33"
 /tissue_type="pooled lung tumors"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: lung; Vector: pRT73D-Pac (Pharmacia) with a
 modified polylinker; Site_1: NotI; Site_2: EcoRI; 1st
 strand cDNA was prepared from mRNA obtained from pooled
 lung tumors with a Not I - oligo(dT) primer [5',
 TGTTCACATCTGACGTGAGCGGCGGCTCTGTTTGTGTGTGTGTGTGTGTGTGTGTGT
 3'].
 Double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not
 I and Eco RI sites of the modified pRT73 vector. Library
 went through one round of normalization, and was
 constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 280 a 201 c 292 g 306 t
 ORIGIN
 Query Match 66.1%; Score 753.4; DB 10; Length 1079;

GAGAGAGAGCGCGCCCAATTAATCTCGAGTAAATTAATTAATCCCGCCCC 3'1. cDNA
was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3'
end: SstI. Host: SOLR.

FEATURES

source

Location/Qualifiers

1..722
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="MCD:MGI:1892879"
/db_xref="taxon:10090"
/clone="1810023A07"
/sex="male"
/tissue_type="pancreas"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="10 day old"
1..722
/gene="Rnf7"
misc_feature
1..722
/gene="Rnf7"
/note="data source:MCD, source key:MGI:1337096,
evidence:ISS
ring finger protein 7"
BASE COUNT 191 a 147 c 201 g 183 t
ORIGIN

Query Match 61.5%: Score 700.8; DB 11: Length 722;
Best Local Similarity .99.3%: Pred. No. 1.5e-149;

Matches 715; Conservative 0; Mismatches 2; Indels 3; Gaps 1;

19 GGGCGAGGTGAGAGAGGAGGAGAACCTGCTCTTCTTCGCTCGAGGAGCGCAGG 78
|||||
2 GGGCGAGGTGAGAGAGGAGGAGAACCTGCTCTTCTTCGCTCGAGGAGCGCAGG 61
|||||
79 CTCCAGTGGGAGGAGGAGCAAGATGTTCTCTCAAGAAAGTGAACCGGATGCTG 138
|||||
62 CTCCAGTGGGAGGAGGAGCAAGATGTTCTCTCAAGAAAGTGAACCGGATGCTG 121
|||||
139 GAGCTGGAGAGTGAAGGAGTACCTGTGCAATCTGAGGAGTCCAGGTGATGATGCTG 198
|||||
122 GAGCTGGAGAGTGAAGGAGTACCTGTGCAATCTGAGGAGTCCAGGTGATGATGCTG 181
|||||
199 CCTTCAGTCAAGCTGAAGAAACAGAGAGAGAGTGTGTGTGGGAGAGTGTAA 258
|||||
182 CTTTCAGTCAAGCTGAAGAAACAGAGAGAGAGTGTGTGTGGGAGAGTGTAA 241
|||||
259 CCATTCTTCACAACTGCTGCAATGCTGCTGTGGTGAAGACAAATCGCTGCTCT 318
|||||
242 CCATTCTTCACAACTGCTGCAATGCTGCTGTGGTGAAGACAAATCGCTGCTCT 301
|||||
319 GTCCAGAGAGAGTGGGTGTGTGTCCAAAGAAATCGCAANTGAGAGTGGCCAGGCGCTCG 378
|||||
302 GTCCAGAGAGAGTGGGTGTGTGTCCAAAGAAATCGCAANTGAGAGTGGCCAGGCGCTCG 361
|||||
379 GGTGTGTGTGCTGACCTGTGACCAAGAGTAAACACTGACAGGGATTCATCTTGAAGAG 438
|||||
362 GGTGTGTGTGCTGACCTGTGACCAAGAGTAAACACTGACAGGGATTCATCTTGAAGAG 421
|||||
439 AAGAGATGCTGTGCGCTTTGAGACTACCAAAAGCTTGTATTAATTTGTCTGTTTA 498
|||||
422 AAGAGATGCTGTGCGCTTTGAGACTACCAAAAGCTTGTATTAATTTGTCTGTTTA 481
|||||
499 GTTGTGGAAATCTCTACAAATTAAGATTAATTTGTTAAATAGGCTTTCCTACCTG 558
|||||
482 GTTGTGGAAATCTCTACAAATTAAGATTAATTTGTTAAATAGGCTTTCCTACCTG 541
|||||
559 TGTGTGTGTGTGATGCAATGATAGAGAGAGCAACACAGAAATATGATCTTGTGTTA 618
|||||
542 TGTGTGTGTGTGATGCAATGATAGAGAGAGCAACACAGAAATATGATCTTGTGTTA 601
|||||
619 TGTGTGTGTGTGATGCAATGATAGAGAGAGCAACACAGAAATATGATCTTGTGTTA 601
|||||
602 TGTGTGTGTGTGATGCAATGATAGAGAGAGCAACACAGAAATATGATCTTGTGTTA 601
|||||

QY 676 TGAGGGTTAAATAATGATTAACGAATGTTACAGTAACAATAATGATTGATTAAGCC 735
|||||
Db 662 TGAGGGTTAAATAATGATTAACGAATGTTACAGTAACAATAATGATTGATTAAGCC 721
|||||

RESULT 8

B1647378 731 bp mRNA linear EST 12-SEP-2001
LOCUS 603279634F1 NIH_CGAP_Man3 Mus musculus cDNA clone IMAGE:5319872 5',
DEFINITION mRNA sequence.
ACCESSION B1647378.1 GI:15561614
VERSION B1647378
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 731)
NIH-MGC <http://mgi.mcl.nih.gov/>.

AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)

TITLE Unpublished (1999)

JOURNAL Contact: Robert Strausberg, Ph.D.

COMMENT Email: cgapbs-remail.nih.gov

Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: L1AM1811 row: c column: 09

High quality sequence stop: 731.

FEATURES

source

Location/Qualifiers

1..731
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:5319872"
/clone_lib="NIH_CGAP_Man3"
/tissue_type="tumor, gross tissue"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-Sport6; Site:1: NotI;
Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert 2 kb. Library constructed by Life
Technologies, catalog #12017-018. Investigators providing
samples: Lothar Hennighausen/Chu-Xia Deng, NIH Reference
for transgenic model: Xu et al., Nature Genetics 22, 37-43
(1999). Note: this is a MGI_CGAP Library."
BASE COUNT 192 a 157 c 200 g 182 t
ORIGIN

Query Match 60.2%: Score 686.6; DB 10: Length 731;

Best Local Similarity 98.8%: Pred. No. 2.6e-146;

Matches 724; Conservative 0; Mismatches 4; Indels 5; Gaps 3;

21 CGAGCTGAGAGAGCGCGGAGAACCTGGTCTTCTTCGCACTCGGAGCGAGGCT 80
|||||
Db 1 CGAGCTGAGAGAGCGCGGAGAACCTGGTCTTCTTCGCACTCGGAGCGAGGCT 60
|||||
QY 81 CCAAGTCGGGAGCGCAAGATGTTCTCTCAAGAAAGTGAACGCGGTAGCCATGTGA 140
|||||
Db 61 CCAAGTCGGGAGCGCAAGATGTTCTCTCAAGAAAGTGAACGCGGTAGCCATGTGA 120
|||||
QY 141 GCTGGAGAGCTTGAGTCCGATCTGTGCGCATCTGCAAGGCTCCAGGTGATGATGCTGCC 200
|||||
Db 121 GCTGGAGAGCTTGAGTCCGATCTGTGCGCATCTGCAAGGCTCCAGGTGATGATGCTGCC 180
|||||
QY 201 TTGATGTCAAGTGAAGAAACAGAGAGAGTGTGTGGTGTGGGAGAGATGTAAC 260
|||||
Db 181 TTGATGTCAAGTGAAGAAACAGAGAGAGTGTGTGGTGTGGGAGAGATGTAAC 240
|||||
QY 261 ATTCTTCCACAACCTGCTGATGCTCCCTGTGGTGAAGAGAAACAATCGCTGCTCTGT 320
|||||

Db 241 ATTCCTTCACACGCTGCTGATGTCCTGTGTGGTGAAGACAGAACATCGCTGCCCTGT 300
 QY 321 GCCAGCAGAGATGGGTAGTCCAAAGAAATCGGCAATGAGAGTGGCCAGGCGCTCTCG 380
 Db 301 GCCAGCAGAGATGGGTAGTCCAAAGAAATCGGCAATGAGAGTGGCCAGGCGCTCTCG 360
 QY 381 TGTGTTGCTGACCTGGACAAAGACTAAACACTGACAGGGATTCATCTTGAAGAGAG 440
 Db 361 TGTGTTGCTGACCTGGACAAAGACTAAACACTGACAGGGATTCATCTTGAAGAGAG 420
 QY 441 AGAATGCTGTGCGCTTTGAGACTACCAAGCGCTGCTTTATTAATTTGCTGTAGT 500
 Db 421 AGAATGCTGTGCGCTTTGAGACTACCAAGCGCTGCTTTATTAATTTGCTGTAGT 480
 QY 501 TTTGGAAATCTCTACAAATTAAGTATTTGTTAAAAATGGCTTCCTCTCTGCTG 560
 Db 481 TTTGGAAATCTCTACAAATTAAGTATTTGTTAAAAATGGCTTCCTCTCTGCTG 540
 QY 561 TGTGTGTGTATACGAATGATAGAGAGCGAAGACACAGAAAATGATCTTTGTTATC 620
 Db 541 TGTGTGTGTATACGAATGATAGAGAGCGAAGACACAGAAAATGATCTTTGTTATC 600
 QY 621 TGTACCCAGCACT---GGAATCTGTGTACAGAGACACTGTGTTTGTATGCTG 677
 Db 601 TGTACCCAGCACTGTGGAACATTTGTGTACAGAGACCA-TGTCACTGTATATGCTG 659
 QY 678 AGGTTTAAAAATAGATTAACGAATGTATACATTAACAATTAATGATTAAGAGCGA 737
 Db 660 AGGTTTAAAAATAGATTAACGAATGTATACATTAACAATTAATGATTAAGAGCGA 718
 QY 738 CTCCTCTAATCC 750
 Db 719 CTCCTCTAATCC 731

RESULT 9
 B1156184 978 bp mRNA linear EST 05-JUL-2001
 LOCUS 602903365F1 NIH_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5032984.5,
 mRNA sequence.
 ACCESSION B1156184
 VERSION B1156184.1 GI:14616185
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NIH-MGC http://mgc.ncl.nih.gov/
 1 (bases 1 to 978)
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: rgs@bbs-rtmail.nih.gov
 Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LINL at:
 http://image.lnl.gov
 Plate: L14M1092 row: a column: 17
 High quality sequence stop: 855.
 Location/Qualifiers
 1..978
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="IMAGE:5032984"
 /clone_id="NIH_CGAP_Mam3"
 /tissue_type="tumor, gross tissue"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: pCMV-SPORT6; site:1: NotI;
 site:2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert 2 Kb. Library constructed by Life

Technologies, catalog #12017-018. Investigators providing
 samples: Lothar Hennighausen/Chu-Xia Deng, NIH Reference
 for transgenic model: Xu et al., Nature Genetics 22, 37-43
 (1999). Note: this is a NCBI_CGAP Library."

BASE COUNT 250 a 199 c 260 g 269 t
 ORIGIN

Query Match 59.1%; Score 673.4; DB 10; Length 978;
 Best Local Similarity 90.5%; Pred. No. 2.8e-143;
 Matches 837; Conservative 0; Mismatches 61; Indels 27; Gaps 10;

QY 75 CAGGCTCCAAAGTGGAGGAGCAAGATGTTCTCTCAAGAAAGTGAGACGCGTAGCCA 134
 Db 1 CAGGCTCCAAAGTGGAGGAGCAAGATGTTCTCTCAAGAAAGTGAGACGCGTAGCCA 60
 QY 135 TGTGAGACTGGAGCTGATGCGATACCTGTGCGCAATCTCAGAGTCCAGGTATGAGT 194
 Db 61 TGTGAGACTGGAGCTGATGCGATACCTGTGCGCAATCTCAGAGTCCAGGTATGAGT 120
 QY 195 CCTGCTTCGATGCAAGCTGAAAAACAAGCAAGAGCACTGTGTGCTGTGGGAGAGT 254
 Db 121 CCTGCTTCGATGCAAGCTGAAAAACAAGCAAGAGCACTGTGTGCTGTGGGAGAGT 180
 QY 255 GTAACCAATTCCTTCACACACTGCTGCATGTCCTGTGCGTGAACACAGCAATCCTGCC 314
 Db 181 GTAACCAATTCCTTCACACACTGCTGCATGTCCTGTGCGTGAACACAGCAATCCTGCC 240
 QY 315 CTCGTGTCGACGAGGAGCTGGGTAGTCCAAAGAAATGCGCAATGAGAGTGGCCAGGCG 374
 Db 241 CTCGTGTCGACGAGGAGCTGGGTAGTCCAAAGAAATGCGCAATGAGAGTGGCCAGGCG 300
 QY 375 TCCGTGTGTGTGCTGCTACCCCTGGACAAAGACTAAACACTGCAAGGGATTCATCCTTGAG 434
 Db 301 TCCGTGTGTGTGCTGCTACCCCTGGACAAAGACTAAACACTGCAAGGGATTCATCCTTGAG 360
 QY 435 AGAGAGAGATGCTGTGCGCTTTGAGACTACCAAGCGCTGCTTTATTAATTTGCTG 494
 Db 361 AGAGAGAGATGCTGTGCGCTTTGAGACTACCAAGCGCTGCTTTATTAATTTGCTG 420
 QY 495 TTTATGTTGGGAATCTCTACAAATTAAGTATTTGTTAAAAATGGCTTCCTCACT 554
 Db 421 TTTATGTTGGGAATCTCTACAAATTAAGTATTTGTTAAAAATGGCTTCCTCACT 480
 QY 555 CTGCTGTGTGTGTGTATACGAATGATAGAGAGCGAAGACACAGAAAATGATCTTGG 614
 Db 481 CTGCTGTGTGTGTGTATACGAATGATAGAGAGCGAAGACACAGAAAATGATCTTGG 540
 QY 615 TTTATGCTACCCAGCACT---GGAATGCTGTTCACAGAGAACTGTTGTGTTA 671
 Db 541 TTTATGCTACCCAGCACTGTGGAACATTTGTTCACAGAGAACTGTTGTGTTA 600
 QY 672 TGTGTGAGAGGTTAAAAATAGATTAACGAATG--TTACAGTAACAATAATGATGATGAA 730
 Db 601 TGTGTGAGAGGTTAAAAATAGATTAACGAATGTTTACAGTAACAATAATGATGATGAA 660
 QY 731 AAGCGAATCTCTCTAATCTTTTGTGTGGAGAGAGCAAGCGAG-----GCCACC 785
 Db 661 AAGCGAATCTCTCTAATCTTTTGTGTGGAGAGAGCAAGCGAGGCCACCCTT 720
 QY 786 TGTGCTCTTATTTGCT-----GTGAATGAGAGATTTTAACTGCACTCAGTAAG-AGGC 839
 Db 721 TGTGCTCTTATTTGCTTTGTAAGAGCGATTTTAACTGCACTCAGTAAGAGAGCC 780
 QY 840 GTAACCTGCGGG--TAACGTATATATGCGCTA---CTGTGCGGTAACGCGCTTGTGTC 893
 Db 781 GTAACCTGCGGGTAACCTGTAATATGCGCTAACCTGCGGTAACGCGCTTGTGTC 840
 QY 894 T-CCTGACTCTCCATCTTT---GACTTGGCCAGAGAGCC-TGGATTTGTTCAACACTT 947
 Db 841 TCCCTGACTCTCCATCTTTGAACTTTGGGCCAGAGAGCTTGGATTTGTTCAACACTT 900
 QY 948 AGTTCTAAAGACATGTTTCTGTTT 972

/note="Organ: liver; Vector: pcwv-SF076; Site-1: NotI; Site-2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.9 kb. Constructed by Life Technologies. Note: this is a NCI CGAP Library."

BASE COUNT 198 a 166 c 210 g 180 t

ORIGIN

Query Match 57.1%; Score 651.4; DB 10; Length 754;
Best Local Similarity 96.2%; Pred. No. 2.9e-138;
Matches 711; Conservative 0; Mismatches 21; Indels 7; Gaps 4;

20 GCCGACGTGAGAGAGCGGAGAGACCCCTGCTTTTTCGCACTCCGGAGGCGAGGC 79
1 GCCGACGTGAGAGAGCGGAGAGACCCCTGCTTTTTCGCACTCCGGAGGCGAGGC 60
80 TCCAGTGGGAGGAGGAGACAGATGTTCTCTCAAGAGTGAACGGCGTACCATGTGG 139
61 TCCAGTGGGAGGAGGAGGAGACAGATGTTCTCTCAAGAGTGAACGGCGTACCATGTGG 120
140 AGCTGGAGCTGAGTGGCATACCTGTCATCTGACGGTCCAGGTCAGTGTATGCTGCG 199
121 AGCTGGAGCTGAGTGGCATACCTGTCATCTGACGGTCCAGGTCAGTGTATGCTGCG 180
200 CTTCGATGTCAGCTGAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 259
181 CTTCGATGTCAGCTGAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
260 CATTCCTCCACAACTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 319
241 CATTCCTCCACAACTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
320 TGGCAGCAGAGCTGGGTAGTCCAAAGATGCGCAATGAGAGTGGCCCGAGGCGCTCTG 379
301 TGGCAGCAGAGCTGGGTAGTCCAAAGATGCGCAATGAGAGTGGCCCGAGGCGCTCTG 360
380 GT 439
361 GT 420
440 GAGATGCTGTGGCCCTTTTGAAGATCAGCAAGAGCTGCTTATTAATTGCTGTG 499
421 GAGATGCTGTGGCCCTTTTGAAGATCAGCAAGAGCTGCTTATTAATTGCTGTG 480
500 TTTTGGGAATTCCTCAATTAAGAT--AATTTGTTAAATGGCCTTCCACTCTG 557
481 TTTTGGGAATTCCTCAATTAAGAT--AATTTGTTAAATGGCCTTCCACTCTG 540
558 GT 617
541 GT 600
618 ATCTGTACCCAGCAGCTGTGGGACATTTGTGTACAGAGAACTGTGCTTAATGTC 674
601 ATCTGTACCCAGCAGCTGTGGGACATTTGTGTGTACAGAGAACTGTGCTTAATGTC 660
675 TTGGAGGTTA-AAAAATAGATTAACGAA-TGTTACAGTAAACAATAAATGATTGAAA 732
661 TTGGAGGTTAACAATATGATTAACGAAATTTGTAACAGTAAACAATAAATGCTCGAACA 720
733 GCCGACCTCCTCTAATCT 751
721 GCCGACCTCCTCTAATCT 739

RESULT 12
LOCUS BFO18804/c 704 bp mRNA - linear EST-29-DEC-2000
DEFINITION ux3b03.x1 McCarrey Eddy type B spermatogonia Mus musculus cDNA
clone IMAGE:3655085.3, similar to TR:09W121.09W121 ZINC RING-FINGER
PROTEIN SAG. - mRNA sequence.

ACCESSION BFO18804
VERSION BFO18804.1 GI:10750210

KEYWORDS

EST.
house mouse.
Mus musculus

SOURCE

Mus musculus

ORGANISM

Mus musculus

REFERENCE

1 (bases 1 to 704)
Marras, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,
Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person
B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schur, R., Ritter
, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,
Waterson, R. and Wilson, R.
The WashU-NCI Mouse EST Project 1999

AUTHORS

Unpublished (1999)
Other ESTs: ux83b03.y1
Contact: Marra M/MashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LBNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:1417389
Seq primer: Primer name ambiguous
High quality sequence stop: 469.

FEATURES

source

1..704

/organism="Mus musculus"

/strain="CD-1"

/db_xref="taxon:10090"

/clone="IMAGE:3655085"

/clone_lib="McCarrey Eddy type B spermatogonia"

/sex="male"

/tissue_type="type B spermatogonia, pooled from multiple mice"

/dev_stage="8 day"

/lab_host="DH10B (phage-resistant)"

/note="Organ: testis; Vector: pBluescript SK+ (Stratagene

; Site-1: XhoI; Site-2: EcoRI; cDNA oligo dt-primed

[5'-(GA)10-ACTAGTCGAGTTTCTTTTCTTTT-3'] and directionally

cloned using 5' linkers 5'-AATTGGCAGCAG-3' and

5'-CTGCGCCG-3'. Size selection of >400bp material gives

average insert size ranging from 1-2 kb. Library was mass

excised (from lambda-unizap-XR) and resulting

single-stranded phagemids were prepped and transformed

into DH10B. Library contains 968 recombinants.

References: J. Androl. 20:635-639 and Gene 25:263-269.

Library constructed and donated by J. McCarrey, Ph.D. of

(Southwest Foundation for Biomedical Research, Dept. of

Genetics); excision done by E.M. Eddy, Ph.D. (National

Institute of Health, National Institute of Environmental

Health Sciences). Original lambda-based library is

available through ATCC catalog #63417."

BASE COUNT 178 a 185 c 140 g 199 t 2 others

ORIGIN

Query Match 57.1%; Score 651; DB 10; Length 704;
Best Local Similarity 96.4%; Pred. No. 3.5e-138;
Matches 678; Conservative 0; Mismatches 7; Indels 4; Gaps 2;

53 CTTTCTGCACTCCGGAGGAGGCTCCAAAGTCGGAGGCGAGCAAGATGTTCTCTC 112
704 CTTTCTGCACTCCGGAGGAGGCTCCAAAGTCGGAGGCGAGCAAGATGTTCTCTC 645
113 AAGAAGTGAACGCGGTAG-CCATGTGAGCTGGGAGCTGTGAGTGCATCTGTCAT 171
644 AAGAAGTGAACGCGGTAG-CCATGTGAGCTGGGAGCTGTGAGTGCATCTGTCAT 585
172 CTTGAGGCTGCAAGGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 231
584 CTTGAGGCTGCAAGGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 525

B1556918
 LOCUS 737 bp mRNA linear EST 05-SEP-2001
 DEFINITION 603239970F1 NIH_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5292832 5',
 mRNA sequence.
 ACCESSION B1556918
 VERSION B1556918.1 GI:15444232
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 737)
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC).
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L1AM11740 row: 1 column: 17
 High quality sequence stop: 685.
 Location/Qualifiers
 1..737
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="IMAGE:5292832"
 /clone_1lb="NIH_CGAP_Mam3"
 /tissue_type="tumor, gross tissue"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site:1: NotI;
 Site:2: SalI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert 2 kb. Library constructed by Life
 Technologies, catalog #12017-018. Investigators providing
 samples: Lothar Hennighausen/Chu-Xia Deng, NIH Reference
 for transgenic model: Xu et al., Nature Genetics 22, 37-43
 (1999). Note: this is a NCI_CGAP Library."
 BASE COUNT 200 a 152 c 212 g 173 t
 ORIGIN
 Query Match 56.7%; Score 646.2; DB 10; Length 737;
 Best Local Similarity 97.1%; Pred. No. 4.4e-137;
 Matches 702; Conservative 0; Mismatches 13; Indels 8; Gaps 4;

QY 380 GTGTGCTGCTGACCTGGACAAAGACTAAACATGACAGGGATTCATCTTGAGAGAGA 439
 |||||
 Db 361 GATGTGCTGCTGACCTTGACAAAGACTAAACATGACAGGGATTCATCTTGAGAGAGA 420
 QY 440 GAGGATGCTGTGCGCCTTTGAGACTCACCAAGCGCTTGTATTTATTTGCTGTTTG 499
 |||||
 Db 421 GAGGATGCTGTGCGCCTTTGAGACTCACCAAGCGCTTGTATTTATTTGCTGTTTG 480
 QY 500 TTTTGGAAATTCCTACAAATTAAGATTAATTTGT-AAAAAGGCGCTTCTACCTG 558
 |||||
 Db 481 TTTGGGAAATTCCTACAAATTAAGATTAATTTGT-AAAAAGGCGCTTCTACCTG 540
 QY 559 TGTGTGTGTGTATACGAATGCATAGAGAGCGAGAACCCAGAAATGATCTTTGTTTA 618
 |||||
 Db 541 TGTGTGTGTGTATACGAATGCATAGAGAGCGAGAACCCAGAAATGATCTTTGTTTA 600
 QY 619 TCTGTACCCAGACT--GGAACATTGTGTTACAGAGACATTTTGTGTTATGCT 675
 |||||
 Db 601 TCTGTACCCAGACTGTGGACATGATGTTCACAGAGAACATG--AGGTAATGCT 657
 QY 676 TGAGGGTTAAATATGATTAACGAATGTACGTAACAATTAAT-TCATTGAAAAC 734
 |||||
 Db 658 TGAGGGTTAAATATGATTAACGAATGTACGTAACAATTAATTCATTGAAAAC 717
 QY 735 CGA 737
 |||
 Db 718 CGA 720

RESULT 15
 BF302267 822 bp mRNA linear EST 21-NOV-2000
 LOCUS 602031233F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:416288 5',
 DEFINITION mRNA sequence.
 ACCESSION BF302267
 VERSION BF302267.1 GI:11248790
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 822)
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L1AM9455 row: a column: 09
 High quality sequence stop: 738.
 Location/Qualifiers
 1..822
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:416288"
 /clone_1lb="NCI_CGAP_SG2"
 /lab_host="DH10B (V1 phage-resistant)"
 /note="Organ: salivary gland; Vector: pCMV-SPORT6; Site:1:
 NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo
 dt. Average insert size 1.3 kb. Constructed by Life
 Technologies. Note: this is a NCI_CGAP Library."
 BASE COUNT 227 a 178 c 229 g 188 t
 ORIGIN

Query Match 54.8%; Score 624.4; DB 10; Length 822;

Best Local Similarity 96.9%; Pred. No. 4.2e-132;
Matches 713; Conservative 0; Mismatches 11; Indels 12; Gaps 7;

```
OY 6 GGGCGCGCGCCCATGGCCGACGCGAGAGAGGAGAGAACCTGCGCTCTTCTTCGCACT 65
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5 GGGTGCGCGCCCATGGCCGACGCGAGAGAGGAGAGAACCTGCGCTCTTCTTCGCACT 64
OY 66 CCGGAGCGCAGGCTCCAAAGTGGGAGGAGACAGATGTTCTCTCAAGAGTGAACG 125
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 65 CCGGAGCGCAGGCTCCAAAGTGGGAGGAGACAGATGTTCTCTCAAGAGTGAACG 124
OY 126 CCGTAGCCATGTGAGCTGGGAGCTTGAAGTGGATACCTGTGCACTGCGAGGTCAGG 185
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 125 CCGTAGCCATGTGAGCTGGGAGCTTGAAGTGGATACCTGTGCACTGCGAGGTCAGG 184
OY 186 TGATGGATGCCCTGCTTGATGTCAAGCTGAAGAAACAGACAGAGACTGTGTGTGCT 245
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 185 TGATGGATGCCCTGCTTGATGTCAAGCTGAAGAAACAGACAGAGACTGTGTGTGCT 244
OY 246 GGGGAGAGTGTAAACCATTCCTTCCAAACTGCTGCATGTCCTGTGGTGAAACAGACA 305
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 245 GGGGAGAGTGTAAACCATTCCTTCCAAACTGCTGCATGTCCTGTGGTGAAACAGACA 304
OY 306 ATCGCTGCCCTGTGTGCCAGCAGACTGGGTAGTCCAAAGAAATCGGCAAAATGAGAGTGG 365
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 305 ATCGCTGCCCTGTGTGCCAGCAGACTGGGTAGTCCAAAGAAATCGGCAAAATGAGAGTGG 364
OY 366 CCCAGGCGCTCCTGTGTGGTGTGCTGACCTGGACAAAGACTAAACACTGCGAGGGATTG 425
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 365 CCCAGGCGCTCCTGTGTGGTGTGCTGACCTGGACAAAGACTAAACACTGCGAGGGATTG 424
OY 426 ATCCTTGAGAGAGAGAGAGATGCTGTGCGCTTGAGACTCACCAAGGCTTCTTATTA 485
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 425 ATCCTTGAGAGAGAGAGATGCTGTGCGCTTGAGACTCA--AAAGCTTCTTATTA 482
OY 486 ATTGTCTGTGTTAGTTTGGGAAATTCCTACATTAAGATAATTTGTTAAAAATGGCCT 545
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 483 AT--TGCTGTAGTTGGGAAATTCCTACATTAAGATAA-TTGTAAAAAATGGCCT 538
OY 546 TTCCACCTCTGCTGTGTGTGTGTATGATGATGATGATGATGATGATGATGATGATGAT 605
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 539 TTCCACCTCTGCTGTGTGTGTGTGTATGATGATGATGATGATGATGATGATGATGAT 598
OY 606 TGATCTTTGTTATCTGTACCCAGCACT--GGAACATTTGTTACAGAGAGACATTTGT 662
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 599 TGATCTTTGTTATCTGTACCCAGCACTGTGGAACTTTGTTACAGAGAGACA-TGT 657
OY 663 TTGTGTTTATGCTTGAGGGTT-AAAAAATAGATTAACGAATGTACAGTAAACAATAAAA 721
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 658 TTGTGTTTATGCTTGAGGGTTCAACAATAAGATAAACGAATGTACAGT-ACAAATTA 716
OY 722 TGCATTGAAAAGCCGA 737
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 717 TGCATTGAAAAGCCGA 732
```

Search completed: July 25, 2002, 18:05:39
Job time: 14218 sec

THIS PAGE BLANK (USPTO)